Perception audiovisuelle de la parole chez le sourd postlingual implanté cochléaire et le sujet normo-entendant : étude longitudinale psychophysique et neurofonctionnelle

Julien Rouger

▶ **To cite this version:**

Julien Rouger. Perception audiovisuelle de la parole chez le sourd postlingual implanté cochléaire et le sujet normo-entendant : étude longitudinale psychophysique et neurofonctionnelle. Neurosciences [q-bio.NC]. Université Paul Sabatier - Toulouse III, 2007. Français. NNT : . tel-00364405

HAL Id: tel-00364405

https://theses.hal.science/tel-00364405

Submitted on 26 Feb 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Toulouse III – Paul Sabatier
U.F.R. Sciences de la Vie et de la Terre

Thèse

pour obtenir le grade de
Docteur de l'Université de Toulouse
délivré par l'Université Toulouse III – Paul Sabatier
Discipline : Neurosciences

présentée et soutenue publiquement par

Julien Rouger
le 29 novembre 2007

Perception audiovisuelle de la parole chez le sourd postlingual implanté cochléaire et le sujet normo-entendant : étude longitudinale psychophysique et neurofonctionnelle

Jury

Pascal Barone	Directeur de thèse
Olivier Deguine	Co-directeur de thèse
Christian Lorenzi	Rapporteur
Jean-Luc Schwartz	Rapporteur
Lionel Collet	Examinateur
Pier-Giorgio Zanone	Président du jury

Ce mémoire de thèse a été rédigé sous LaTeX, un langage de composition typographique spécifiquement adapté à la production de documents scientifiques et mathématiques.

Les outils nécessaires à l'utilisation de LaTeX sont facilement accessibles sur Internet sous forme de logiciels libres gratuits ; par ailleurs leur mise à jour est continuellement assurée par une vaste communauté de programmeurs. Ici le compilateur et l'éditeur utilisés sont respectivement MiKTeX, disponible sur le site http://www.miktex.org, et LaTeXEditor, disponible sur le site http://www.latexeditor.org.

Les références bibliographiques sont gérées par une base de données au format BibTeX. Actuellement il n'existe pas encore de programmes de conversion des références au format EndNote ou PubMed vers le format BibTeX. Pour les utilisateurs intéressés j'ai créé deux utilitaires informatiques qui permettent de gérer automatiquement ces conversions. Ces programmes sont disponibles sur demande (j.rouger@psy.gla.ac.uk).

De très nombreux documents de présentation et forums d'aide sur LaTeX et BibTeX sont disponibles sur Internet. En guise d'introduction on pourra par exemple consulter *"Une courte (?) introduction à LaTeX2e"* à l'adresse : http://www.laas.fr/~matthieu/cours/latex2e/.

Remerciements

Je tiens tout d'abord à remercier du fond du cœur tous les patients implantés cochléaires qui nous ont fait confiance et ont accepté de nous prêter une oreille attentive, ainsi que les nombreux volontaires normo-entendants qui ont bien voulu nous accorder un peu de leur temps.

Je remercie Michèle Fabre-Thorpe, Pascal Barone et Yves Trotter pour m'avoir accueilli au sein du laboratoire Cerveau et Cognition UMR CNRS 5549, ainsi que Jean-Michel Lassalle pour son accueil au sein de l'Ecole Doctorale CLESCO de l'Université Paul Sabatier. Merci à Olivier Deguine et Bernard Fraysse pour nous avoir ouvert les portes du service d'oto-rhino-laryngologie du CHU Purpan, à Marie-Laurence Laborde pour nous avoir reçu dans l'unité Implant Auditif de réhabilitation orthophonique, ainsi qu'à Jean-François Démonet pour son accueil dans l'unité Inserm 825.

Ma gratitude s'adresse tout particulièrement à Pascal Barone pour avoir assuré pendant ces quatre années la direction et l'encadrement de mon travail de thèse, pour sa présence constante et pour sa patience devant ma détermination à progresser à mon rythme, lentement mais sûrement. Elle s'adresse également à Olivier Deguine pour avoir assuré la co-direction de cette thèse, pour avoir particulièrement facilité notre insertion dans l'univers hospitalier, pour sa rigueur clinique ainsi que pour son profond sens humain. Elle s'adresse enfin à Jean-François Démonet pour son rôle indispensable dans la définition de nos protocoles de neuroimagerie et pour son aide dans l'interprétation linguistique de nos résultats comportementaux et neurofonctionnels.

Je remercie Christian Lorenzi et Jean-Luc Schwartz pour avoir accepté d'être rapporteurs de cette thèse, ainsi que Lionel Collet et Pier-Giorgio Zanone pour avoir rejoint le jury en tant qu'examinateurs. Merci au comité ACI Neurosciences Intégratives et Computationnelles (n° NIC 0085, juin 2003) et à la Fondation pour la Recherche Médicale pour avoir retenu la candidature de notre projet et pour nous avoir permis de le financer du début jusqu'à la fin. Merci enfin au comité CCPPRB de Toulouse pour nous avoir autorisé à mettre en place le protocole TEP.

Je suis particulièrement reconnaissant envers Sébastien Lagleyre pour sa contribution essentielle à ce travail et pour ses qualités de dynamisme, d'enthousiasme et d'organisation, de même qu'envers Elie Eter, David Sauvajon, Zoé Massida et Kuzma Strelnikov. Je leur adresse ici des remerciements chaleureux pour nos fructueuses collaborations.

Je tiens à remercier tous les chercheurs, ingénieurs, docteurs et étudiants qui ont contribué à nous aider : Gérard Viaillard, Hélène Gros et tout le personnel du centre TEP du CHU Purpan ; Xavier de Boissezon, Mélanie Jucla et Olivier Dufor de l'unité Inserm 825 ; Sophie Denève de l'ENS d'Ulm ; Chris James de Cochlear et Michel Beliaeff de MedEl ; Daniel Pressnitzer, Jean-Michel Hupé et Renaud Lestringant du CerCo ; Pascal Despretz du CHU de Lille.

Parmi les personnels du CerCo, je remercie également Maxime Rosito, Marc Macé, Nadège Bacon-Macé et Jérémy Fourny pour nous avoir fait partager leurs compétences en programmation informatique, en électronique et en mécanique. Un grand merci à Ca-

therine Marlot pour ses talents de documentaliste émérite, à Claire Thoueilles pour ses compétences logistiques, ainsi qu'à Carmen Olmos et Joël Ferrer pour leurs compétences informatiques !

Merci à tous les étudiants du CerCo pour leur accueil et pour tous ces moments de détente bienvenue partagés ensemble ; merci en particulier à Sébastien, Olivier, Ludovic, Mathieu, Marianne, Carole, Aurélie, Jinsong et Fabrice. Merci également à tous les étudiants de l'Inserm pour leur sympathique accueil.

Je remercie enfin tous les élèves de collège-lycée de Toulouse et de Lille qui ont réussi à me supporter comme prof particulier de maths et de physique, ainsi que tous les élèves de PCEM1 qui ont eu la patience de suivre mes cours de psychomotricité à la fac de médecine de Rangueil.

Pour terminer ces remerciements, je tiens à exprimer ma profonde gratitude envers les personnes qui m'ont constamment apporté leur soutien, en particulier dans les moments délicats. Je remercie en premier lieu ma mère Ingrid, ma soeur Astrid et mon frère Adrien pour leur écoute et leur patience, ainsi que mon père Pierre qui n'est plus de ce monde mais veille encore sur nous. Merci également à Marie-Martine Cailler et Pierre Fenart pour leur aide spirituelle et leur clairvoyance.

Je remercie tous mes proches et mes compagnons de route pour leur présence et pour tout ce que nous avons pu partager : Myriam, Christine, Etienne et tous les jeunes et moins jeunes du mouvement des Focolari pour leur dynamisme œcouménique ; Blandine et l'agence *Mission Impossible* pour m'avoir fait découvrir les trésors cachés de la Picardie ; Claire pour son accueil au foyer Saint-Dominique et pour son originalité créatrice ; Nana et les grenouillettes du foyer pour leur bonne humeur et leur joie de vivre ; Magali, Caroline, Michel & Lucile, Guillaume, Mathieu et la communauté Foi & Lumière pour tous ces précieux instants de chaleureuse humanité ; Waël, Jean-Paul et le cercle des pianistes disparus pour ces harmonies célestes ; Isabelle et tous ses randonneurs en herbe ; Laurent, Mathias, Claire & Alex, Audrey & Chris pour leurs cocktails détonnants de psychologie et d'ingénierie ; Sébastien & Laetitia, Leïla & Nicolas pour toutes ces escapades physiques et métaphysiques, ainsi que Jean-Pierre Nicola et Richard Pellard pour avoir repoussé les limites de la connaissance humaine.

Merci à la « Forestière » pour son délicieux moelleux au chocolat !

Merci à César Franck et Frédéric Chopin pour leurs compositions intemporelles.

Merci à toutes celles et ceux que j'ai oublié, qu'ils me pardonnent !

Table des matières

Introduction générale **11**
Organisation du mémoire . . . 13

I Audition, parole et perception multisensorielle **15**

1 L'audition **17**
1.1 Système auditif périphérique . . . 18
1.2 Perception auditive . . . 26
1.3 Système auditif central . . . 29
Bibliographie . . . 35

2 Production de la parole **41**
2.1 Appareil phonatoire . . . 42
2.2 Aspects articulatoires de la phonation . . . 46
2.3 Caractéristiques acoustiques des phonèmes . . . 52
Bibliographie . . . 57

3 Perception audiovisuelle de la parole **59**
3.1 Développement de l'intégration audiovisuelle de la parole . . . 60
3.2 Facilitations perceptives par la parole audiovisuelle . . . 61
3.3 Facteurs influant sur l'intégration audiovisuelle de la parole . . . 62
3.4 Conséquences d'une incongruence spatio-temporelle . . . 64
3.5 Conséquences d'une incongruence phonétique : effet McGurk . . . 65
3.6 Mécanismes d'intégration audiovisuelle de la parole . . . 67
Bibliographie . . . 73

4 Intégration neuronale de la parole et du langage **79**
4.1 Parole et langage . . . 80
4.2 Substrats neuronaux de la parole auditive . . . 81
4.3 Substrats neuronaux de la parole visuelle . . . 82
4.4 Réseau cérébral audio-visuo-moteur de la parole . . . 84
4.5 Réseau cérébral de traitement du langage . . . 91
Bibliographie . . . 97

5 Intégration multisensorielle **109**
5.1 Multisensorialité : principes et propriétés . . . 110
5.2 Multisensorialité et congruence spatio-temporelle . . . 110
5.3 Autres formes d'interaction multisensorielle . . . 113

5.4 Substrats neuronaux de la perception multisensorielle . . . 114
Bibliographie . . . 119

II Surdité, implant cochléaire et réorganisation corticale 125

6 La surdité 127
6.1 Qu'est-ce que la surdité ? . . . 128
6.2 Dépistage de la surdité . . . 129
6.3 Causes de la surdité . . . 134
6.4 Effets de la surdité sur la perception de la parole . . . 137
6.5 Prothèses auditives . . . 138
6.6 Communication avec les personnes sourdes . . . 140
Bibliographie . . . 143

7 L'implant cochléaire 145
7.1 Principe . . . 145
7.2 Indications et critères d'implantation . . . 147
7.3 Implantation cochléaire . . . 148
7.4 Rééducation fonctionnelle . . . 149
7.5 Aspects techniques . . . 149
Bibliographie . . . 153

8 Réorganisation neurofonctionnelle après une altération sensorielle 155
8.1 Propriétés de la neuroplasticité . . . 156
8.2 Neuroplasticité intramodale . . . 161
8.3 Neuroplasticité intermodale après déprivation auditive . . . 171
8.4 Neuroplasticité intermodale après déprivation visuelle . . . 173
Bibliographie . . . 177

9 Réorganisation neurofonctionnelle après implantation cochléaire 191
9.1 Effets de l'implantation sur la perception de la parole . . . 192
9.2 Récupération perceptive : facteurs techniques et chirurgicaux . . . 194
9.3 Récupération perceptive : facteurs neurofonctionnels . . . 198
9.4 Réorganisation corticale après implantation cochléaire . . . 201
Bibliographie . . . 205

III Etude des conséquences de l'implantation cochléaire 215

10 Problématique et méthodes 217
10.1 Problématique . . . 217
10.2 Préparation des stimuli comportementaux . . . 218
10.3 Traitement phonétique des réponses comportementales . . . 224
10.4 Neuroimagerie fonctionnelle . . . 227
Bibliographie . . . 229

11 Etudes comportementales chez le sujet sain **231**
11.1 Lecture labiale chez le sujet sain . . . 232
11.2 Sujet sain : parole masquée par du bruit . . . 235
11.3 Sujet sain : parole traitée par l'implant . . . 242
11.4 Entraînement perceptif avec une simulation d'implant . . . 249
Bibliographie . . . 255

12 Etudes comportementales chez le patient implanté **257**
12.1 Perception de la parole chez le sujet implanté . . . 258
12.2 Intégration audiovisuelle chez le sujet implanté . . . 262
12.3 Lecture labiale chez le sujet implanté . . . 268
12.4 Effet McGurk chez le sujet implanté . . . 271
Bibliographie . . . 285

13 Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral **289**
13.1 Matériel et méthodes . . . 291
13.2 Activations cérébrales chez les patients implantés . . . 294
13.3 Activations cérébrales chez les sujets contrôles . . . 308
13.4 Différences d'activité patients/contrôles . . . 312
13.5 Etude de corrélation des activations cérébrales . . . 319
13.6 Discussion . . . 320
Bibliographie . . . 329

Conclusion générale **333**

Annexes **337**

A Neuroanatomie **339**
A.1 Description anatomique . . . 339
A.2 Normalisation de l'anatomie corticale . . . 347
Bibliographie . . . 349

B Publications **351**
Rouger et al. (2007). Proc. Natl. Acad. Sci. U.S.A. 104, 7295–7300. . . . 353
Rouger et al. (2008). Brain Res. 1188, 87–99. . . . 359
Strelnikov et al. (2009). Neuropsychologia 47, 972–979. . . . 373

Index **381**

« On lui amène un sourd, qui de plus parlait difficilement, et on le supplie de lui imposer les mains. Le prenant à part, loin de la foule, Jésus lui mit les doigts dans les oreilles et lui toucha la langue avec sa propre salive ; puis, levant les yeux au ciel, il soupira et dit : "Ephphata"*, c'est-à-dire* "Ouvre-toi"*. Aussitôt ses oreilles s'ouvrirent, sa langue se délia, et il parlait correctement. »*

— Saint Marc, VII 32–35

« La relation du cerveau à la pensée est donc complexe et subtile. Si vous me demandiez de l'exprimer dans une formule simple, nécessairement grossière, je dirais que le cerveau est un organe de pantomime, et de pantomime seulement. Son rôle est de mimer la vie de l'esprit, de mimer aussi les situations extérieures auxquelles l'esprit doit s'adapter. L'activité cérébrale est à l'activité mentale ce que les mouvements du bâton du chef d'orchestre sont à la symphonie. La symphonie dépasse de tous côtés les mouvements qui la scandent ; la vie de l'esprit déborde de même la vie cérébrale. Mais le cerveau, justement parce qu'il extrait de la vie de l'esprit tout ce qu'elle a de jouable en mouvement et de matérialisable, justement parce qu'il constitue ainsi le point d'insertion de l'esprit dans la matière, assure à tout instant l'adaptation de l'esprit aux circonstances, maintient sans cesse l'esprit en contact avec des réalités. Il n'est donc pas, à proprement parler, organe de pensée, ni de sentiment, ni de conscience ; mais il fait que conscience, sentiment et pensée restent tendus sur la vie réelle et par conséquent capables d'action efficace. Disons, si vous voulez, que le cerveau est l'organe de l'attention à la vie*. »*

— Henri Bergson, L'âme et le corps.
in *L'énergie spirituelle*, ch. II, conférence du 28 avril 1912.

Introduction générale

Parmi les éléments qui distinguent l'être humain de ses cousins animaux, outre ses dispositions pour le rire et l'humour, son sens de l'artistique, du spirituel et de l'abstrait, et parfois sa profonde bêtise, l'une des spécificités essentielles de notre espèce réside dans le développement et l'utilisation particulière que nous faisons du langage parlé. De nombreuses espèces animales utilisent des moyens de communication vocaux, comme par exemple les singes, les baleines, les dauphins ou les oiseaux ; toutefois aucune d'entre elles ne possède la complexité grammaticale ou la capacité à représenter des notions abstraites que l'on peut rencontrer dans le langage humain.

Malgré l'importance du langage écrit, autre spécificité humaine, ou de la communication non verbale véhiculée par les gestes, les expressions du visage ou les intonations de la voix, la communication parlée demeure un facteur essentiel dans la structuration psychologique et sociale des individus. Ce fait devient particulièrement flagrant dans le cas des personnes atteintes d'une surdité partielle ou totale, qui souffrent en premier lieu de la perte de leur capacité à communiquer correctement avec leurs semblables. Par ailleurs, la surdité modifie profondément la nature de l'interaction sociale avec les autres puisque la personne sourde doit concentrer toute son attention visuelle sur les mouvements des lèvres, du visage et des mains de son interlocuteur pour parvenir à saisir le sens de la conversation. Alors que l'audition rendait possible une grande souplesse entre les interlocuteurs, en permettant notamment au regard de l'auditeur de se détourner ailleurs, la surdité impose de sévères limitations à la communication interindividuelle.

Chez l'être humain comme chez certaines espèces animales, l'apprentissage de la communication parlée repose sur deux mécanismes. Tout d'abord, il n'est pas possible de dissocier la voix de l'audition, dans la mesure où l'apparition des vocalisations au cours du développement de l'enfant se fait grâce à une incessante interaction entre la perception des vocalisations des parents et proches, les essais de production articulatoire et la rétroaction auditive qui en résulte. Ainsi le développement de la voix (et donc de la communication parlée) est conditionné par le bon fonctionnement à la fois de l'appareil auditif et de l'appareil phonatoire : c'est la célèbre « boucle audio-phonatoire » mise en avant par Alfred Tomatis. Cela explique qu'un enfant né sourd se trouvera dans l'incapacité d'apprendre à parler malgré la préservation fonctionnelle de son appareil phonatoire : il deviendra ce que l'on appelle un « sourd-muet », bien que le terme « muet » soit employé de façon abusive puisqu'il n'indique pas une déficience des structures articulatoires. De même, un enfant devenu sourd lors de la période d'apprentissage du langage conservera certains défauts de prononciation ainsi qu'une légère altération du timbre de la voix, en raison de la dégradation du retour auditif nécessaire à une acquisition correcte de l'articulation de la parole.

Ensuite, contrairement à ce que l'on a tendance à penser, la parole ne se restreint pas au domaine de l'audition : en effet la parole possède également une nature visuelle de par les mouvements visibles des lèvres, de la langue et du visage qui accompagnent les fluctuations temporelles et spectrales de la voix. Ainsi la parole n'est pas un signal purement auditif mais un signal audiovisuel à part entière, et l'apprentissage de la parole au cours de l'enfance se fait par intégration simultanée de ses composantes auditive

(fluctuations de la voix) et visuelle (mouvements du visage). En conséquence, la parole est perçue comme un tout audiovisuel et non pas comme un flux auditif accessoirement complété par des informations visuelles. Ce fait est particulièrement bien mis en évidence par l'effet McGurk, qui montre que la perception de la parole est bien le résultat d'une intégration perceptive audiovisuelle et non pas d'une perception purement auditive. La nature fondamentalement audiovisuelle de la parole est également illustrée dans le cas des enfants nés aveugles, qui rencontrent des difficultés pour apprendre à distinguer certains phonèmes proches sur le plan auditif mais distincts sur le plan visuel.

De façon plus générale, la parole est un cas particulier d'activité multisensorielle. En effet, la plupart des éléments qui composent notre environnement physique sont des sources d'informations à la fois visuelles et auditives, voire également tactiles ou même olfactives et gustatives. Il existe donc une cohérence spatiale et temporelle d'origine physique entre les différents canaux sensoriels par lesquels nous percevons le monde qui nous entoure. L'étude du comportement humain et animal ainsi que l'analyse des activations neuronales associées indiquent que la nature a su tirer parti de cette cohérence physique en permettant, dès les plus bas niveaux de traitement cérébral, une intégration multisensorielle des signaux en provenance des différentes modalités sensorielles.

Un objectif commun à tous les âges de l'humanité est de pouvoir suppléer aux défaillances physiques du corps humain, que celles-ci trouvent leur origine dans le vieillissement naturel des organes, dans des affections génétiques ou encore dans des agressions par des agents biologiques, mécaniques, chimiques ou radiatifs. Ce besoin de « réparer » le corps humain est particulièrement évident dans les cas où la défaillance constatée empêche les organes indispensables à la survie d'assurer convenablement leurs fonctions ou prive l'être humain d'une partie de sa capacité à interagir avec son environnement, mais également dans les cas où cette défaillance, même mineure d'un point de vue fonctionnel, est une source de souffrance physique ou psychologique.

Au cours du XX^{e} siècle, les avancées scientifiques dans la connaissance de la structure et du fonctionnement du corps humain, dans la production des matériaux bio-compatibles et dans la miniaturisation électronique ont rendu possible des applications prothétiques difficilement imaginables pour les générations des siècles précédents : on dispose aujourd'hui de dispositifs de substitution qui assurent les mêmes fonctions physiologiques que le cœur, les reins ou le foie et les mêmes fonctions mécaniques que les articulations de la hanche ou du genou. Toutefois il est encore difficile d'égaler l'extrême complexité du corps humain au niveau sensoriel. Dans ce domaine, les progrès technologiques les plus remarquables ont été accomplis pour l'audition, en particulier avec l'invention des implants cochléaires.

Un implant cochléaire est une neuroprothèse se substituant aux premiers étages fonctionnels du système auditif, incluant l'oreille externe, l'oreille moyenne et une grande partie de l'oreille interne. L'implant cochléaire, indiqué principalement en cas de surdité bilatérale sévère à profonde, a pour rôle de capter les sons de l'environnement, de les convertir en impulsions électriques puis de stimuler directement les fibres du nerf auditif. Ces implusions électriques artificielles sont alors acheminées vers les centres nerveux, où elles seront traitées pour aboutir à une sensation auditive.

Les premiers implants cochléaires étaient assez rudimentaires et ne permettaient de retransmettre qu'un seul canal de fréquences sonores. Aujourd'hui certains implants

peuvent stimuler jusqu'à 22 canaux de fréquences, ce qui est encore très loin des milliers de fréquences que l'appareil auditif humain est susceptible de discriminer. Par ailleurs on se restreint presque systématiquement à une implantation unilatérale en raison de critères cliniques, les effets de l'implantation bilatérale étant pour l'instant peu connus, ainsi que de critères otologiques mais également économiques. Les données de la littérature scientifique actuelle ainsi que la progression constante des connaissances technologiques permettent d'imaginer que sur le long terme, l'usage tendra à privilégier les implantations bilatérales avec des résolutions atteignant plusieurs dizaines de canaux de fréquences.

Une personne atteinte de surdité postlinguale (après l'acquisition du langage parlé) assiste à la disparition soudaine ou progressive de la perception auditive et se voit donc obligée de reporter son attention sur la modalité visuelle. Les études scientifiques menées sur des personnes sourdes révèlent de multiples réorganisations dans les mécanismes perceptifs fonctionnels et corticaux, en particulier dans ceux intervenant lors de la perception de la parole. Après une implantation cochléaire, la stimulation électrique du nerf auditif par l'implant restaure une certaine forme de sensation auditive, avec toutefois une qualité bien moindre que pour un système auditif intact.

Cette réintroduction de la sensation auditive après une période de surdité soulève plusieurs questions : quelles vont en être les conséquences au niveau fonctionnel et cortical ? Comment cette nouvelle forme d'audition va-t-elle être intégrée avec la vision de la parole, qui elle n'a pas été modifiée ? Comment les capacités de perception de la parole vont-elles évoluer sur le long terme, aussi bien dans des conditions visuelle ou auditive que dans un contexte audiovisuel ? Assisterons-nous à un phénomène de réorganisation des réseaux cérébraux impliqués dans la perception de la parole ? Le cas échéant, comment cette réorganisation plastique va-t-elle évoluer et y aura t-il une différenciation de l'activité cérébrale par rapport à des personnes normo-entendantes ?

Nous tenterons dans la suite d'apporter des éléments de réponse à ces différentes problématiques, au travers d'une approche comportementale et neurofonctionnelle de la perception audiovisuelle de la parole chez les patients sourds postlinguaux implantés cochléaires. Nous complèterons cette approche en effectuant de façon aussi systématique que possible des comparaisons avec le cas des sujets normo-entendants pour des conditions perceptives similaires.

Organisation du mémoire

L'un de nos principaux objectifs dans la rédaction de ce mémoire a été de fournir au lecteur une base concise et synthétique des champs de connaissance relatifs à notre étude. Par souci de clarté, nous avons donc choisi de subdiviser ce mémoire en trois parties ; les deux premières dressent un bilan introductif des connaissances disponibles dans la littérature scientifique chez le sujet normo-entendant (*« Audition, parole et perception multisensorielle »*, part. I, p. 17) puis chez le sujet sourd implanté cochléaire (*« Surdité, implant cochléaire et réorganisation corticale »*, part. II, p. 127), tandis que la troisième et dernière décrit l'ensemble des études expérimentales que nous avons mené à terme (*« Etude des conséquences de l'implantation cochléaire »*, part. III, p. 217).

Notre première partie vise à résumer les processus connus intervenant chez le sujet normo-entendant dans l'audition (ch. 1, p. 17), la production de la parole (ch. 2, p. 41), la perception de la parole (ch. 3, p. 59) et la perception du langage (ch. 4, p. 79), puis

de façon plus générale la perception multisensorielle (ch. 5, p. 109). Chaque fois que le sujet s'y prêtait nous en avons abordé les aspects neurologiques en effectuant une revue synoptique de la littérature scientifique.

La deuxième partie concerne les thèmes centraux de notre étude : la surdité (ch. 6, p. 127) et l'implantation cochléaire (ch. 7, p. 145). Nous complétons ces thématiques en décrivant dans un premier temps les conséquences neurofonctionnelles d'une déprivation sensorielle, dont la surdité est un cas particulier (ch. 8, p. 155), puis les conséquences neurofonctionnelles d'une implantation cochléaire (ch. 9, p. 191). Nous aborderons en particulier la question de la plasticité corticale, intimement liée aux notions de réorganisation fonctionnelle et corticale.

La troisième partie est consacrée aux recherches que nous avons effectuées ; elle met notamment en parallèle notre double approche psychophysique et neurofonctionnelle de l'implantation cochléaire avec la revue de la littérature dressée dans les deux parties précédentes. Nous avons regroupé ces recherches en trois volets : études comportementales chez le sujet sain (ch. 11, p. 231) et le patient implanté (ch. 12, p. 257), puis études neurofonctionnelles chez l'implanté unilatéral (ch. 13, p. 289).

Première partie

Audition, parole et perception multisensorielle

Chapitre 1

L'audition

Sommaire

1.1 Système auditif périphérique 18
1.1.1 Anatomie du système auditif 18
1.1.2 Fonctionnement de la cochlée 22
1.2 Perception auditive 26
1.2.1 Seuils perceptifs en fréquence et en intensité 26
1.2.2 Niveaux auditifs subjectifs : échelle des phones 27
1.2.3 Filtres auditifs et masquage psychoacoustique 27
1.3 Système auditif central 29
1.3.1 Relais subcorticaux 29
1.3.2 Cortex auditif 31
1.3.3 Anatomie fonctionnelle du cortex auditif 32
Bibliographie 35

Introduction

Les êtres vivants reçoivent des informations sur leur corps et le monde qui les entoure à travers le filtre spécifique de leurs perceptions. Chez l'être humain comme pour de nombreuses espèces animales, on distingue les systèmes sensoriels qui traitent l'information en provenance de l'environnement extérieur, incluant la vision, l'audition, l'olfaction et la gustation, de ceux qui traitent l'information en provenance de l'organisme, incluant la somesthésie et la perception vestibulaire.

La vision et l'audition se complètent pour véhiculer une proportion très importante des informations environnementales. Chez l'être humain, la vision permet de percevoir les radiations électromagnétiques aux longueurs d'onde comprises entre 400 (violet) et 800 nm (rouge), tout en fournissant une excellente sélectivité spatiale mais une précision temporelle limitée. L'audition permet quant à elle de percevoir les vibrations acoustiques aux fréquences comprises entre 20 et 20 000 Hz, avec une bonne résolution temporelle mais une résolution spatiale limitée.

Les informations audiovisuelles vont ensuite être complétées par la somesthésie, qui renseigne sur les perceptions au niveau cutané (extéroception), au niveau musculaire

et articulaire (proprioception) ainsi qu'au niveau viscéral (interoception). La perception vestibulaire fournit pour sa part des indices sur la position, l'orientation et les accélérations de la tête, tandis que l'olfaction et la gustation donnent certaines informations sur les propriétés physico-chimiques des odeurs et des goûts.

Nous nous intéresserons dans la suite aux particularités de l'audition, en étudiant les différents mécanismes physiologiques et neurologiques qui, à partir d'une vibration acoustique provenant de l'environnement sonore, génèrent la sensation auditive correspondante.

1.1 Système auditif périphérique

Le système auditif périphérique† permet de transformer les vibrations acoustiques en provenance de l'environnement en impulsions nerveuses. Celles-ci seront ensuite communiquées au système auditif central, où elles vont être traitées et intégrées afin d'aboutir à une sensation auditive. Lors de cette transformation des sons en influx nerveux, les différentes parties aériennes, liquides et solides qui constituent le système auditif périphérique vont interagir mécaniquement avec les ondes acoustiques, ce qui aura pour effet de moduler la répartition spectrale de l'intensité acoustique.

1.1.1 Anatomie du système auditif[1]

Oreille externe

L'oreille externe† (voir fig. 1.1, p. 19) est formée du pavillon auditif et du conduit auditif externe. Le conduit auditif, d'une longueur moyenne de 25 mm et d'un diamètre moyen de 7 mm, est le prolongement interne de la conque du pavillon. Sa partie la plus externe est cartilagineuse, tandis que sa partie la plus interne est osseuse et fermée par une membrane souple, le tympan†.

Le rôle de l'oreille externe est de capter les sons de l'environnement, de les amplifier et de les acheminer vers le tympan. La configuration géométrique du pavillon et du conduit auditif permet d'assurer une transmission optimale des ondes acoustiques de fréquences moyennes, le conduit auditif servant alors de résonateur acoustique pour ces fréquences particulières, plus significatives pour la perception de l'environnement sonore et tout particulièrement de la parole humaine. Cette géométrie spécifique contribue également à faciliter la localisation des sons par de légères modifications de leurs composantes spectrales en fonction de leur direction de provenance.

Oreille moyenne

L'oreille moyenne† (voir fig. 1.1, p. 19) est formée du tympan et de la caisse du tympan, où se situe la chaîne des osselets. La caisse du tympan est une cavité remplie d'air située dans la région mastoïdienne de l'os temporal. Elle est séparée de l'oreille externe par le tympan et communique avec l'oreille interne par deux ouvertures membraneuses dans la paroi osseuse : la fenêtre ovale† et la fenêtre ronde†. Cette cavité communique avec le pharynx par l'intermédiaire de la trompe d'Eustache, qui est susceptible de s'ouvrir et se refermer (lors de la déglutition par exemple) pour équilibrer les pressions de part et d'autre du tympan. Les fluctuations de pression acoustique au niveau de la membrane tympanique provoquent des vibrations de cette dernière. Ces vibrations mécaniques sont ensuite transmises à l'oreille interne par la chaîne des osselets.

La chaîne des osselets† est composée de trois os très délicats : le marteau†, fixé sur le tympan par son manche, l'enclume† et l'étrier†, dont l'extrêmité est une petite plaque osseuse

[1] d'après Galifret (1991); Pujol *et al.* (1999)

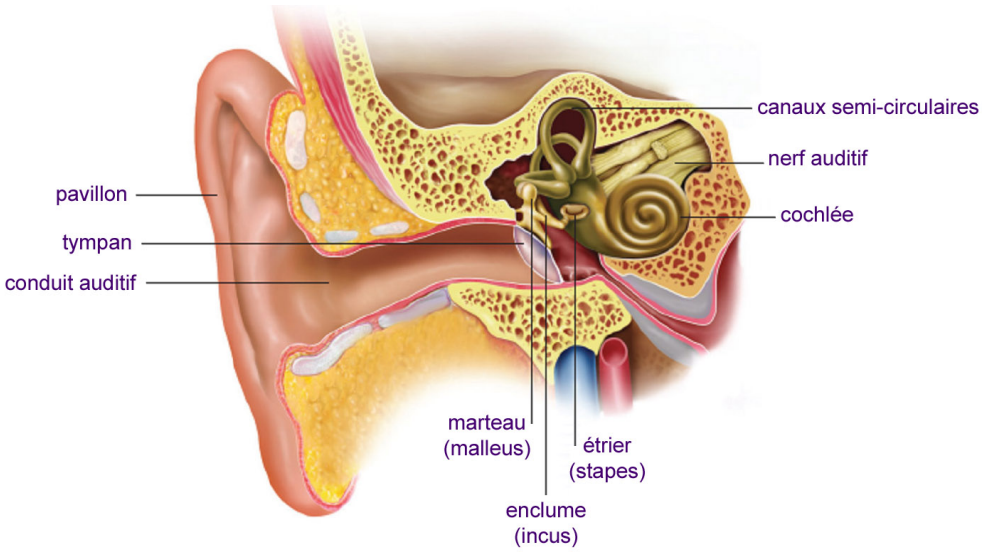


FIG. 1.1 – Schéma anatomique du système auditif

s'insérant dans la fenêtre ovale. Le marteau et l'enclume sont liés par une articulation assez rigide tandis que la liaison de la base de l'enclume avec l'étrier est une rotule très mobile. L'ensemble de ces trois osselets se comporte comme un système de levier susceptible d'amplifier les mouvements vibratoires de la membrane tympanique. L'amplitude de ces mouvements est contrôlée par deux petits muscles : le tenseur du tympan, qui s'insère sur le marteau et dont la contraction réduit la mobilité de la membrane tympanique, et le muscle de l'étrier, ou stapédius, dont la contraction réduit la mobilité de la platine de l'étrier au niveau de la fenêtre ovale.

Le système des osselets de l'oreille moyenne permet de transmettre les vibrations du tympan à la fenêtre ovale, tout en amplifiant et en adaptant ces mouvements mécaniques aux caractéristiques vibratoires de l'oreille interne. Il assure une conversion optimale entre les vibrations acoustiques de l'air du conduit auditif et les vibrations acoustiques des liquides de l'oreille interne. La contraction des muscles des osselets permet en outre d'assurer la protection de l'oreille interne contre les sons de forte intensité. Cette régulation se produit de façon réflexe lors de brusques augmentations de l'intensité acoustique. Elle se produit également lors de la vocalisation, la contraction du stapédius précédant l'émission vocale. Ce contrôle permet de limiter l'impact des sons de la parole sur l'audition de la personne qui émet ces sons.

Oreille interne : cochlée

L'oreille interne† (voir fig. 1.1, p. 19) est constituée par les organes situés dans l'épaisseur du rocher, l'os constituant la paroi interne de la caisse du tympan. Ces organes sont regroupés sous le nom de labyrinthe en raison de leur forme complexe. Le labyrinthe postérieur est l'organe de l'équilibre. Il comprend les trois canaux semi-circulaires, sensibles au déplacement de la tête dans les trois plans de l'espace. Le labyrinthe antérieur, appelé cochlée† (anciennement limaçon), est l'organe de l'audition. La cochlée (voir fig. 1.2, p. 20) se présente sous la forme d'un tube creux tronconique, le tube cochléaire†, enroulé en spirale sur deux tours et demi autour d'un axe creux, le modiolus†.

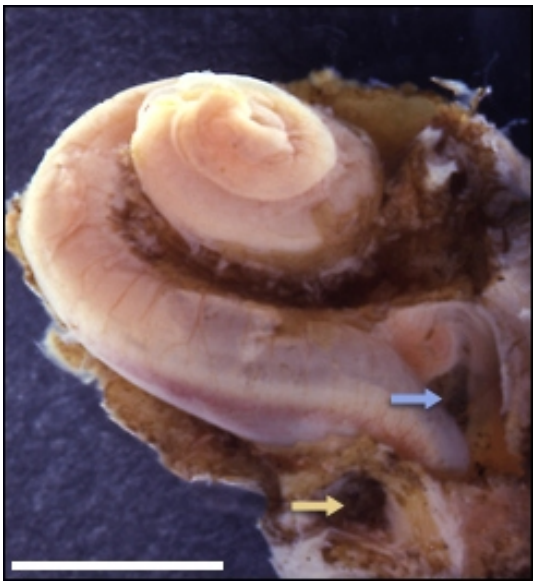


FIG. 1.2 – Cochlée de fœtus humain vue en microscopie électronique. La flèche jaune indique la fenêtre ronde ; la flèche bleue la fenêtre ovale. Echelle : 5 mm. D'après Pujol *et al.* (1999).

Sur une coupe transversale, le tube cochléaire apparaît divisé en trois canaux (voir fig. 1.3, p. 21) : la rampe vestibulaire† (*scala vestibularis*), la rampe tympanique† (*scala tympani*) et la rampe médiane (*scala media*), ou canal cochléaire†. Les rampes vestibulaire et tympanique sont séparées par la lame osseuse spirale à proximité du modiolus et par le canal cochléaire à la périphérie de la cochlée. Le canal cochléaire est séparé de la rampe tympanique par la membrane basilaire† et de la rampe vestibulaire par la membrane de Reissner† (voir fig. 1.4, p. 21).

Au sommet de la cochlée, la lame osseuse spirale et le canal cochléaire s'effacent, et les rampes vestibulaire et tympanique sont mises en communication par un orifice, l'hélicotrème (voir fig. 1.3, p. 21). Les deux rampes sont donc remplies par un même fluide, la périlymphe†. Au contraire, le canal cochléaire ne communique pas avec les autres canaux et est rempli par un fluide, appelé endolymphe†, dont la composition chimique diffère de celle de la périlymphe. À la base de la cochlée, la rampe vestibulaire se termine par la fenêtre ovale, solidaire de la platine de l'étrier, tandis que la rampe tympanique se termine par la fenêtre ronde.

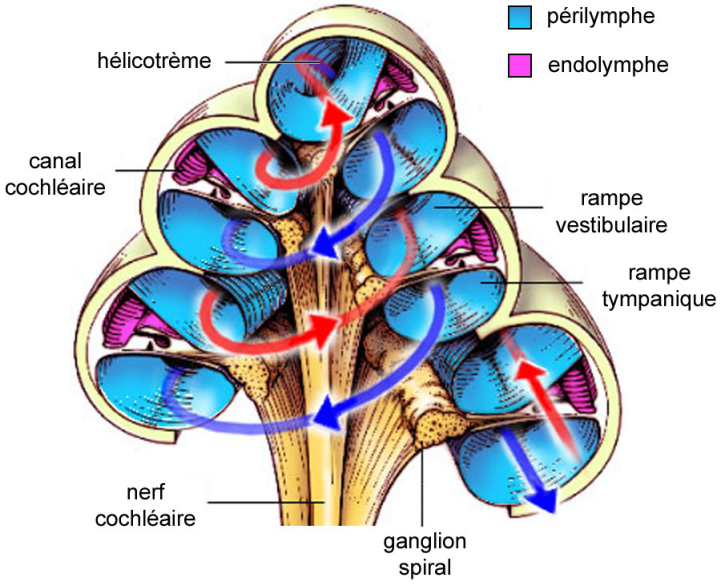


FIG. 1.3 – Section axiale de la cochlée. On peut observer l'enroulement du canal cochléaire contenant l'endolymphe (en violet) et celui des rampes vestibulaire et tympanique contenant la périlymphe (en bleu). La flèche rouge part de la fenêtre ovale et remonte la rampe vestibulaire de la base vers l'apex pour aboutir à l'hélicotrème ; la flèche bleue part de l'hélicotrème et redescend la rampe tympanique de l'apex vers la base pour aboutir à la fenêtre ronde. Adapté d'après Pujol *et al.* (1999).

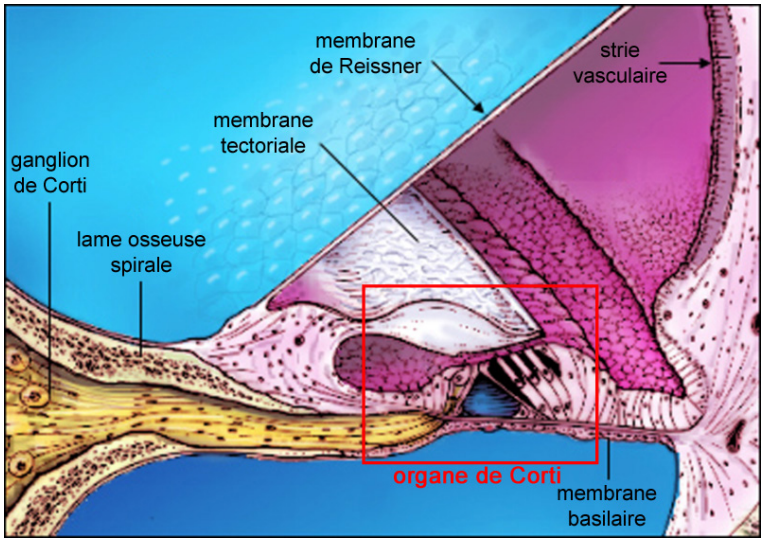


FIG. 1.4 – Section d'un tour de spire de la cochlée. Adapté d'après Pujol *et al.* (1999).

Oreille interne : organe de Corti

La membrane basilaire, qui prolonge la lame osseuse spirale, se fixe à son bord externe à la paroi cochléaire par le ligament spiral, très élastique. Elle supporte l'organe de Corti† (voir fig. 1.5, p. 22 et fig. 1.6, p. 23), qui assure la transduction des vibrations acoustiques des liquides cochléaires en messages nerveux.

L'organe de Corti est constitué de cellules ciliées† disposées en rangées de part et d'autre du tunnel de Corti†, de la base au sommet de la cochlée, chaque rangée comportant une centaine de cellules par millimètre (voir fig. 1.7, p. 23). Parmi les cellules ciliées on distingue les cellules ciliées internes, qui forment une rangée située entre le modiolus et le tunnel de Corti, et les cellules ciliées externes, qui forment trois à quatre rangées situées au-delà du tunnel de Corti (voir fig. 1.5, p. 22 et fig. 1.6, p. 23).

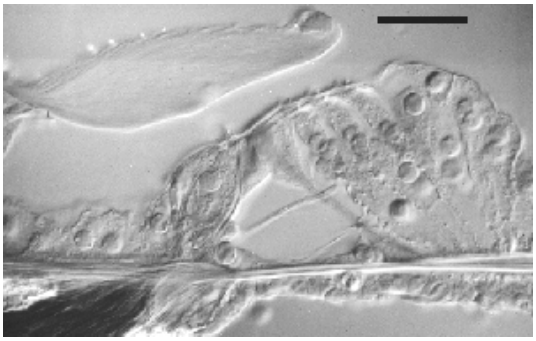


FIG. 1.5 – Section transversale de l'organe de Corti chez le cobaye adulte (microscopie optique Nomarski ; échelle 10 µm) – On peut observer la membrane tectoriale en haut à gauche et la membrane basilaire en bas, horizontalement, sur toute la longueur de la photographie. Au-dessus de celle-ci, de gauche à droite, on distingue l'unique rangée de cellules ciliées internes, puis le tunnel de Corti, puis finalement les rangées de cellules ciliées externes. D'après Pujol *et al.* (1999).

Chaque cellule ciliée présente de nombreux cils, longs de 3 à 6 µm (voir fig. 1.7, p. 23 et fig. 1.8, p. 24). Sur l'organe de Corti repose la membrane tectoriale†, formée d'une mince couche d'un gel glycoprotéique. Les extrêmités des cils des cellules ciliées externes viennent s'insérer dans cette membrane, où elles laissent des empreintes en creux facilement observables, tandis que les cils des cellules ciliées internes y sont reliées par un système assez lâche de minces filaments.

Le nerf auditif compte environ 30 000 fibres nerveuses afférentes. Ces fibres, qui se rassemblent dans le modiolus, sont les axones des neurones dont les corps cellulaires forment les ganglions de Corti, ou ganglions spiraux. Une proportion très importante (environ 90%) de ces neurones innerve les cellules ciliées internes, à raison de quelques dizaines de fibres nerveuses par cellule ciliée. Les fibres nerveuses restantes comportent de nombreuses terminaisons synaptiques qui innervent plusieurs cellules ciliées externes à la fois, suivant ainsi un mode d'innervation complètement différent de celui des cellules ciliées internes. Les neurones innervant les cellules ciliées externes possèdent par ailleurs un corps cellulaire plus réduit.

1.1.2 Fonctionnement de la cochlée[2]

Les mouvements de la platine de l'étrier au niveau de la fenêtre ovale donnent naissance à une onde acoustique dans les liquides cochléaires. La membrane de Reissner, la membrane

[2] d'après Galifret (1991); Pujol *et al.* (1999)

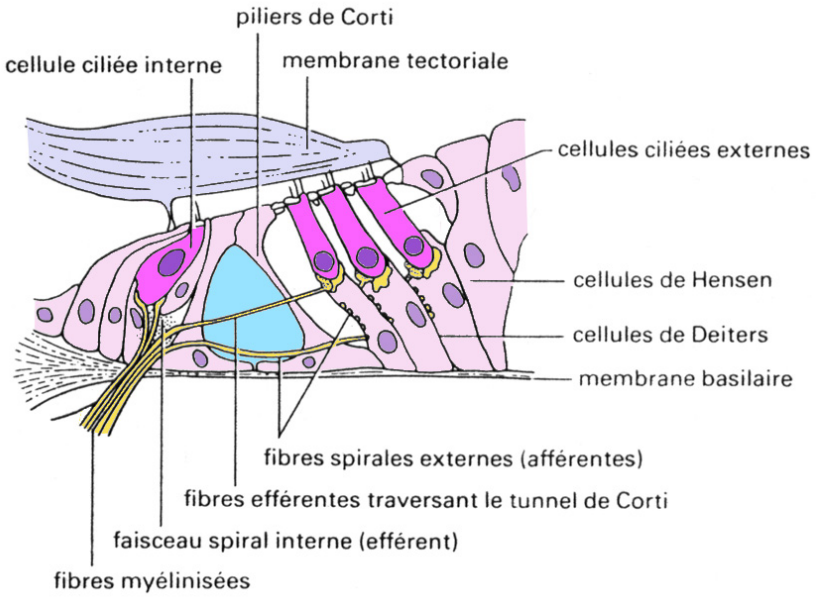


FIG. 1.6 – Schéma de l'organe de Corti vu en coupe transversale. Adapté d'après Galifret (1991).

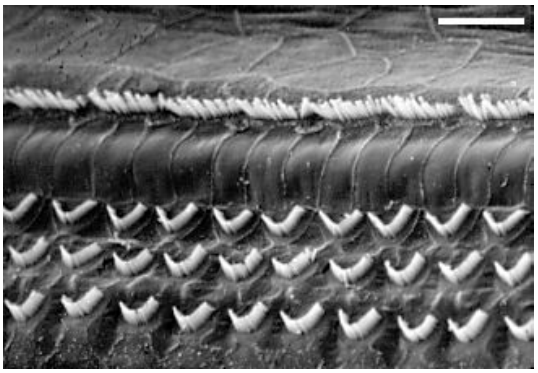


FIG. 1.7 – Vue de dessus de l'organe de Corti chez le rat adulte, à la 5^{e} semaine de développement (microscopie électronique ; échelle 15 µm) – On peut distinguer en haut l'unique rangée de cellules ciliées internes et en bas les trois rangées de cellules ciliées externes. D'après Pujol *et al.* (1999).

basilaire et la membrane de la fenêtre ronde sont alors entraînées par les vibrations acoustiques de ces liquides. Or la membrane basilaire est beaucoup plus rigide à la base qu'au sommet. En effet, alors que le diamètre de la cochlée diminue de la base (2 mm) au sommet (1 mm), la membrane basilaire s'élargit de 100 µm à la base à 500 µm au sommet, tandis que son épaisseur moyenne diminue de la base vers le sommet. Par conséquent, pour une même fréquence de stimulation, l'amplitude des vibrations de la membrane basilaire va dépendre de la position sur la membrane (voir fig. 1.9, p. 24). Il est donc possible d'établir une carte de la position de ce maximum en fonction des fréquences acoustiques, ce qui permet de mettre en évidence une tonotopie cochléaire avec les hautes fréquences codées à la base de la cochlée et les basses fréquences à l'apex (voir fig. 1.10, p. 25). On pourra consulter Moore (1989b) pour une revue de la littérature détaillée concernant la tonotopie cochléaire.

Une fois que la membrane basilaire est entrée en vibration, ses déformations vont se com-

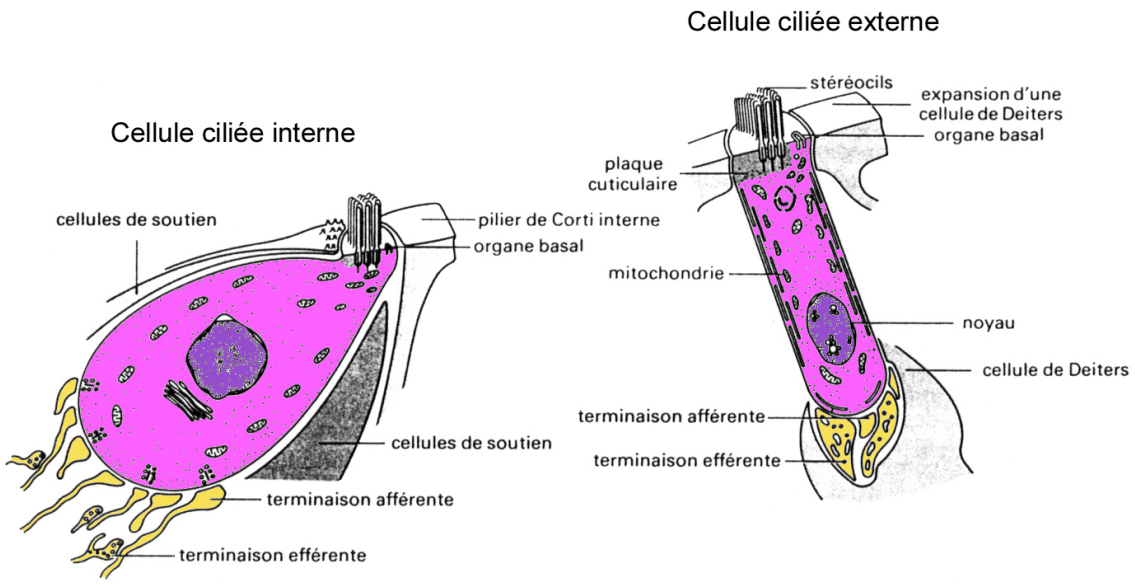


FIG. 1.8 – Structure des cellules ciliées internes et externes. Adapté d'après Galifret (1991).

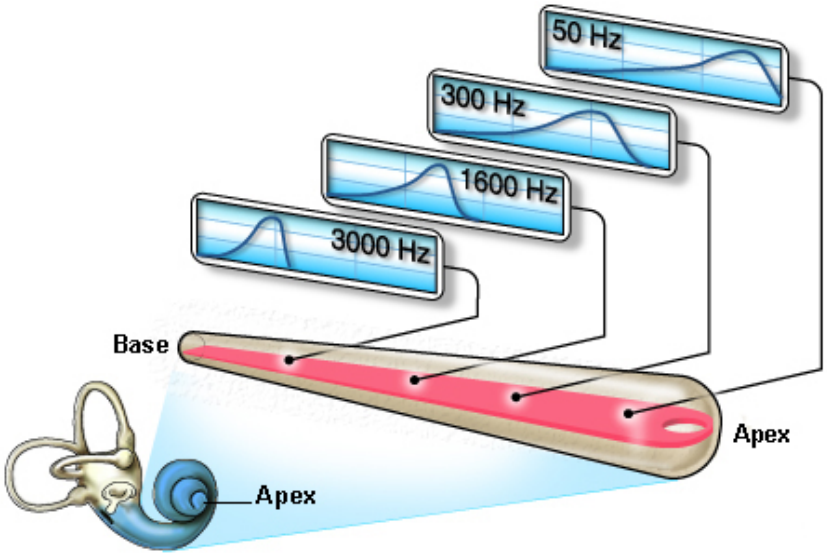


FIG. 1.9 – Répartition de l'amplitude des vibrations de la membrane basilaire en fonction de la fréquence de stimulation. D'après Pujol *et al.* (1999).

muniquer aux cils de l'organe de Corti. Ainsi des déformations de la membrane basilaire vers la rampe vestibulaire vont provoquer l'inclinaison des cils vers l'extérieur de la cochlée tandis que des déformations vers la rampe tympanique vont incliner les cils vers le modiolus (voir fig. 1.11, p. 25). En fonction de l'amplitude, de la fréquence et de la localisation de ces inclinaisons le long de la cochlée, les cellules ciliées vont transmettre des rafales de potentiels d'action aux fibres nerveuses cochléaires correspondantes, ce qui provoquera une stimulation électrique du nerf auditif. Cette stimulation sera alors transmise aux centres nerveux de l'audition, où elle sera traitée et intégrée pour aboutir à une perception auditive.

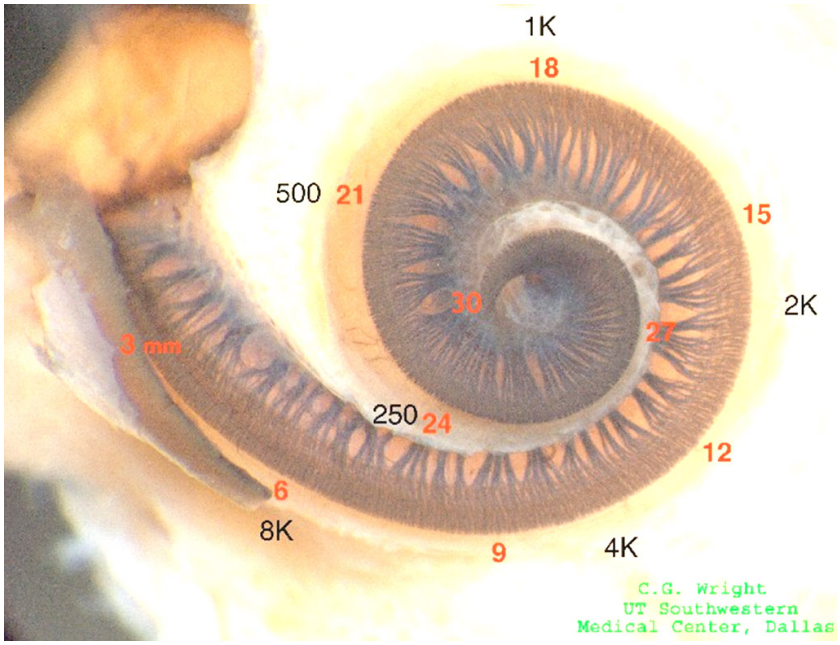


FIG. 1.10 – Tonotopie cochléaire : correspondance entre la profondeur du canal cochléaire (chiffres rouges, en millimètres) et la fréquence sonore associée (chiffres noirs, en hertz et kilohertz). Source : C.G. Wright, UT Southwestern Medical Center, Dallas.

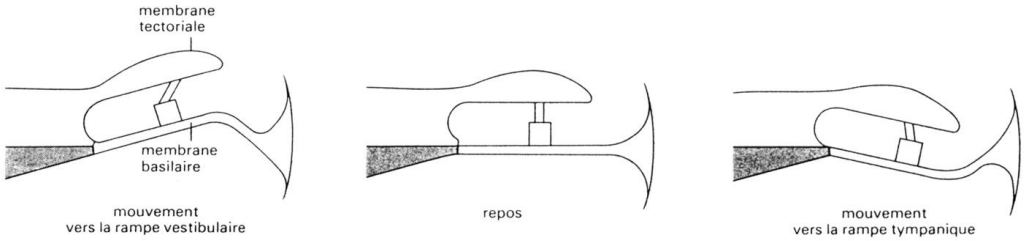


FIG. 1.11 – Mécanismes d'inclinaison des cils lors des oscillations de la membrane basilaire. D'après Galifret (1991).

1.2 Perception auditive

1.2.1 Seuils perceptifs en fréquence et en intensité

Le seuil d'audition normalisé est défini par les limites inférieure et supérieure de sensibilité de l'appareil auditif humain à des stimuli acoustiques, calculées en fonction de la fréquence sur la base d'un très grand nombre de mesures audiométriques réalisées sur des sujets sains. On définit ainsi une courbe de seuil d'audition normal, marquant la limite inférieure moyenne de la sensibilité auditive, et une courbe de seuil intolérable normal, qui en marque la limite supérieure. On appelle aire d'audition la plage auditive délimitée par ces deux seuils.

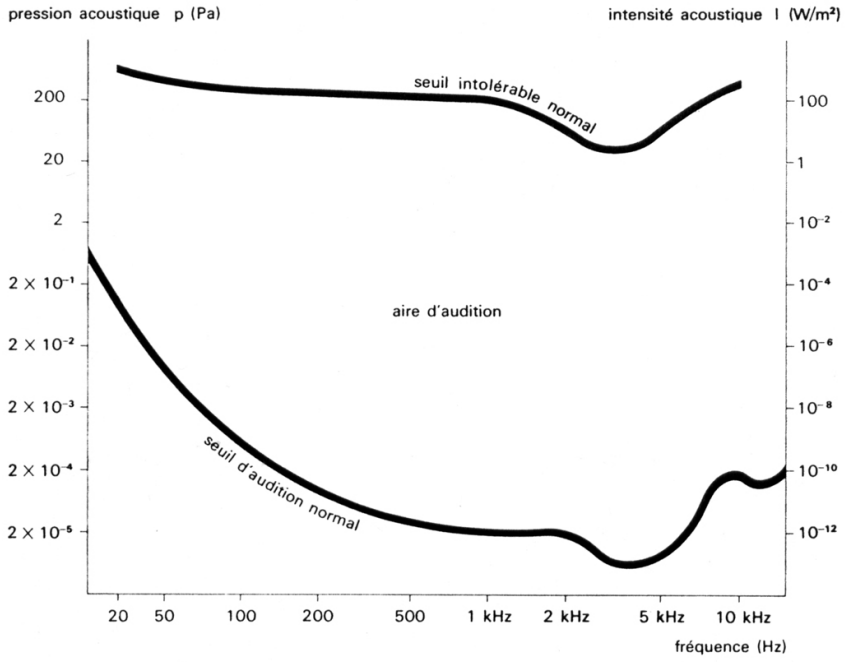


FIG. 1.12 – Courbes des seuils de l'audition humaine. D'après Galifret (1991).

On peut immédiatement remarquer la grande dépendance de la sensibilité du système auditif en fonction de la fréquence. Dans le domaine des fréquences moyennes comprises entre 1000 et 3000 Hz, une intensité sonore de 10^{-12} W.m^{-2} suffit à provoquer une sensation auditive, tandis qu'à une basse fréquence comme 50 Hz, il faut employer une intensité de 10^{-8} W.m^{-2}, 10000 fois plus importante, pour déclencher une perception. La sensibilité du système auditif humain s'échelonne sur une plage d'intensités acoustiques allant approximativement de 10^{-12} à 100 W.m^{-2}, soit un rapport de 10^{14}, et sur une plage de fréquences allant de 20 à 20000 Hz. On appelle infrasons les vibrations de fréquences inférieures à 20 Hz, trop basses pour éveiller une sensation auditive, et ultrasons les vibrations de fréquences supérieures à 20000 Hz, trop élevées pour être perçues par l'audition.

Weber et Fechner ont énoncé une loi permettant de relier de façon très approximative l'intensité acoustique physique et la sensation auditive subjective associée à cette intensité. Cette loi associe une augmentation linéaire de la sensation subjective à une augmentation exponentielle de l'intensité acoustique correspondante. En d'autres termes, nous ressentons la même impression subjective de différence de niveau lorsque l'intensité acoustique varie d'un même facteur :

ainsi, que l'intensité acoustique varie de 1 à 10 ou de 1000 à 10000, nous percevons une même augmentation de niveau sonore. Le niveau subjectif perçu dépend donc de façon logarithmique de l'intensité acoustique. En application de cette loi de dépendance, et au vu du rapport considérable de 10^{14} existant entre les seuils physiques des intensités perceptibles par l'audition, une notation logarithmique a été adoptée pour rendre compte du niveau acoustique des sons.

Ainsi le niveau acoustique d'intensité, exprimé en décibels (dB)†, est défini comme : $N_I = 10\log(\frac{I}{I_0})$, où I est l'intensité acoustique de l'onde et I_0 une intensité acoustique de référence correspondant au seuil de l'audition, fixée à 10^{-12} W.m^{-2}. Une différence de niveau acoustique égale à 20 dB correspond à un rapport d'intensité acoustique égal à 100 ; une différence de 30 dB correspond à un rapport de 1000. On emploie également le niveau acoustique de pression, exprimé aussi en dB (en anglais, dB SPL pour *sound pressure level*), et défini comme : $N_P = 20\log(\frac{P}{P_0})$, où P est la pression acoustique de l'onde et P_0 une pression acoustique de référence correspondant au seuil de l'audition, fixée à 2.10^{-5} Pa. Par comparaison, la pression atmosphérique est près de 10^{10} fois plus importante que cette pression de référence.

1.2.2 Niveaux auditifs subjectifs : échelle des phones

La découverte de la différence de sensibilité du système auditif en fonction de la fréquence des sons a fait apparaître le besoin d'un étalonnage précis de cette sensibilité. Pour deux sons de fréquences différentes et de même intensité acoustique, le niveau perçu est différent. Sans modifier les fréquences initiales, on a alors cherché à introduire une différence d'intensité acoustique dans les sons pour compenser la différence perceptive de niveau et aboutir à des sensations auditives identiques.

Pour cela, on a fixé des niveaux physiologiques de référence en définissant l'échelle perceptive des phones, destinée à caractériser le niveau d'isosonie d'un son. On caractérise un son par un niveau d'isosonie de N phones lorsqu'il est perçu par un auditeur normal moyen à un niveau auditif équivalent à celui perçu pour un son de référence, ce dernier étant défini comme un son pur de niveau d'intensité acoustique N dB et de fréquence 1000 Hz se propageant par ondes planes progressives face à l'auditeur. Cette définition permet d'obtenir des courbes d'égale sensation auditive, ou courbes isosoniques.

On remarque que pour une intensité acoustique donnée, le niveau perçu est beaucoup plus faible pour les basses et hautes fréquences que pour les fréquences moyennes. Cette différence de perception entre les fréquences moyennes et les fréquences extrêmales s'atténue progressivement à mesure que l'intensité acoustique augmente. Une conséquence particulièrement importante est que toute modification du niveau acoustique physique provoque un changement dans le timbre perçu.

1.2.3 Filtres auditifs et masquage psychoacoustique

Les recherches expérimentales en psychoacoustique ont permis de préciser la façon dont le spectre d'un signal sonore incident est traité par le système auditif. Un important ensemble d'études convergent pour indiquer qu'on peut assimiler le traitement acoustique effectué par le système auditif périphérique à une succession de filtres passe-bande se chevauchant, appelés « filtres auditifs† ».

La notion de filtre auditif a été proposée pour la première fois par Fletcher (1940) dans une expérience de masquage psychoacoustique. Le principe de l'expérience d'origine était le suivant : un son pur est masqué par un bruit blanc de bande dont la fréquence centrale est identique à celle du son pur. On fait varier la largeur de bande du bruit blanc tout en conservant la densité

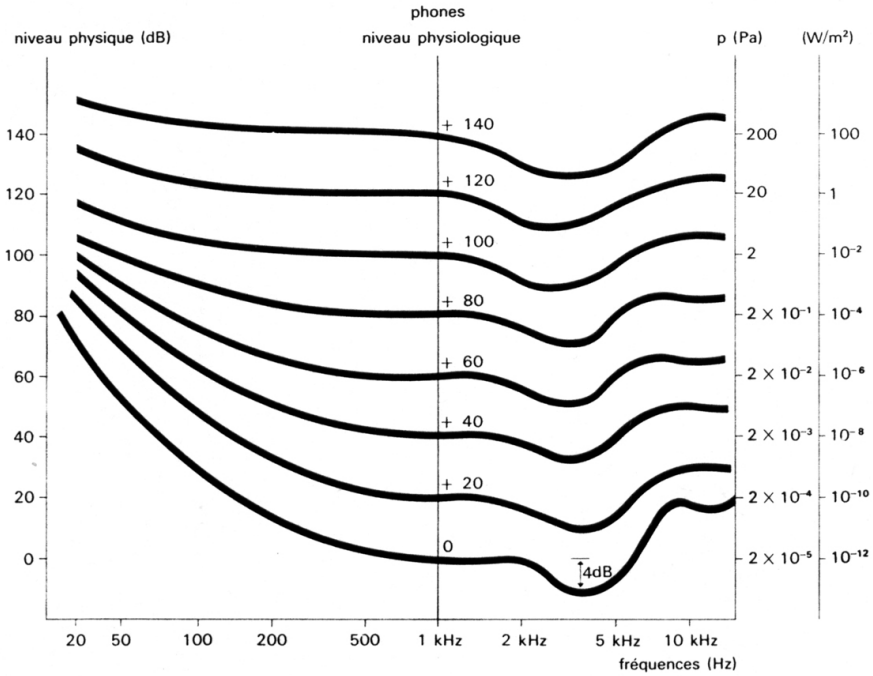


FIG. 1.13 – Courbes isosoniques normalisées de l'audition humaine. D'après Galifret (1991).

Phones	Correspondance	Phones	Correspondance
130	seuil intolérable	50	musique douce
120	réacteur d'avion à 10 m	40	conversation normale
110	atelier industriel	30	résidence tranquille
100	marteau-piqueur à 2 m	20	studio d'enregistrement
90	orchestre symphonique	10	laboratoire d'acoustique
80	rue bruyante	5	chambre muette
60	conversation vive	0	seuil de l'audition

FIG. 1.14 – Echelle de perception des niveaux sonores

spectrale d'énergie constante, de sorte que le niveau sonore augmente avec la largeur de la bande. En exprimant le niveau sonore à partir duquel le son pur est perçu en fonction de la largeur de bande du bruit blanc, on s'aperçoit que ce niveau augmente dans un premier temps, puis se stabilise à partir d'une certaine largeur de bande. Cette *largeur de bande critique* (Fletcher, 1940) correpondrait ainsi à la largeur spectrale d'un filtre auditif passe-bande rectangulaire centré sur la fréquence du son pur utilisé, de sorte que les sons dont la fréquence est située en dehors de cette bande spectrale ne jouent aucun effet masquant pour la fréquence centrale du filtre.

La forme rectangulaire initialement proposée pour les filtres auditifs (Fletcher, 1940) n'est en réalité qu'une approximation grossière. Des méthodes fines permettant de mesurer les formes des filtres auditifs ont été proposées par la suite (Patterson, 1976), ce qui a permis de mettre en évidence une forme quasi-triangulaire dont le sommet correspond à la fréquence centrale. La sélectivité en fréquence des filtres auditifs est toutefois dépendante du niveau sonore de stimulation, et on peut observer l'apparition d'une asymétrie en faveur des basses fréquences lorsque le niveau sonore augmente (Moore, 1989a).

Par ailleurs on peut observer une augmentation de la largeur de bande critique avec la fréquence centrale du filtre auditif. Si l'on caractérise le filtre par sa largeur de bande rectangulaire équivalente (ERB, *equivalent rectangular bandwidth*), qui correspond à la largeur d'un filtre rectangulaire théorique qui laisserait passer la même énergie acoustique totale que le filtre auditif réel, on peut relier l'ERB (expimée en Hz) et la fréquence centrale f du filtre (exprimée en kHz) par la relation suivante : $\text{ERB} = 24.7 \times (4.37f + 1)$ (Moore, 1998). Ainsi chaque ERB correspondrait approximativement à une largeur de 0.9 mm sur la membrane basilaire (Moore, 1998).

1.3 Système auditif central

Le système auditif central† peut être décrit comme une hiérarchie ascendante de différents niveaux de traitement nerveux des informations auditives. Cette hiérarchie part du ganglion spiral pour aboutir au cortex auditif.

1.3.1 Relais subcorticaux

Ganglion spiral Le ganglion spiral† (*spiral ganglion* ; SG) est un relais nerveux situé dans une cavité osseuse en-dessous de l'axe de la cochlée. Il reçoit des projections directes des cellules ciliées de la cochlée et envoie des projections vers les noyaux cochléaires.

Noyaux cochléaires Les noyaux cochléaires† (*cochlear nuclei* ; CN) constituent le premier relais cérébral des signaux nerveux en provenance des nerfs auditifs. Ces noyaux sont situés à l'entrée du tronc cérébral, sur les parties antérieure et postérieure du canal médullaire. Chacun des nerfs auditifs se divise pour se projeter sur les deux noyaux cochléaires. Les noyaux cochléaires sont subdivisés en différentes zones morphologiquement distinctes et montrent des réponses spécifiques aux différentes fréquences auditives. Les basses fréquences sont traitées principalement dans les parties ventrales, rostrales et latérales de chaque noyau, tandis que les hautes fréquences sont traitées dans les parties dorsales, caudales et médiales (Galifret, 1991). On subdivise usuellement chaque noyau cochléaire en une partie antéroventrale (*anteroventral cochlear nucleus* ; AVCN), une partie postéroventrale (*posteroventral cochlear nucleus* ; PVCN) et une partie dorsale (*dorsal cochlear nucleus* ; DCN).

Complexe olivaire supérieur Le complexe olivaire supérieur† (*superior olivary complex* ; SOC) est un groupe de noyaux cérébraux situé également sur le canal médullaire, au-dessus des noyaux cochléaires. Chacune des olives supérieures reçoit des projections en provenance des noyaux cochléaires ipsilatéral et contralatéral. A ce niveau s'effectue donc la première confrontation des messages nerveux en provenance des deux oreilles. Les résultats de la littérature montrent en particulier que les olives supérieures médiales (*medial superior olive* ; MSO) et latérales (*lateral superior olive* ; LSO) jouent un rôle important dans la localisation spatiale des sources sonores (Grothe et Park, 2000; Tollin, 2003). L'olive supérieure médiale est sensible aux différences interaurales de temps d'arrivée des sons (Yin et Chan, 1990), qui permettent de déterminer la direction de la source sonore, tandis que l'olive supérieure latérale traite les différences interaurales d'amplitude entre les sons (Park *et al.*, 1996), qui permettent essentiellement de déterminer l'éloignement de la source sonore.

Noyaux du lemnisque latéral Les noyaux du lemnisque latéral† (*nuclei of the lateral lemniscus* ; NLL) constituent la dernière étape du tronc cérébral. Ils reçoivent des afférences du

complexe olivaire supérieur et des noyaux cochléaires, et envoyant des projections sur le colliculus inférieur. Ils présentent une sensibilité aux indices de phase et d'amplitude.

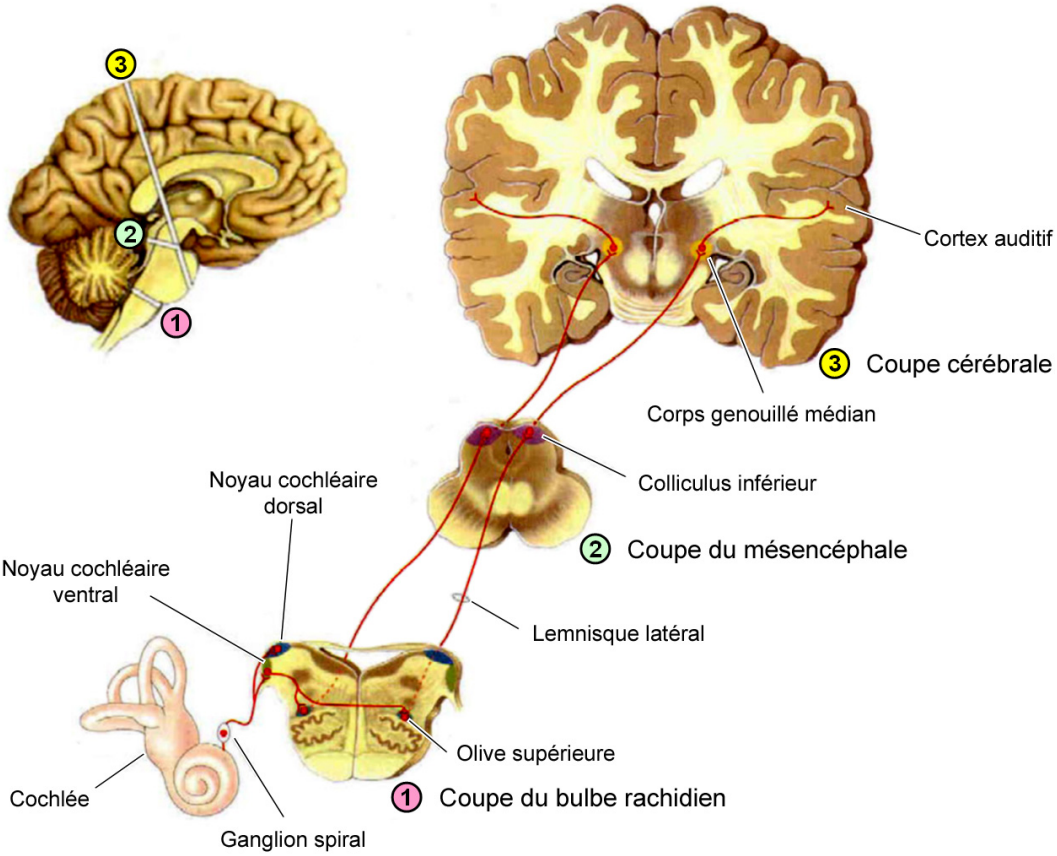


FIG. 1.15 – Schéma d'ensemble des voies nerveuses auditives. Adapté d'après Bear *et al.* (1997).

Colliculus inférieur Le colliculus inférieur† (*inferior colliculus* ; IC) fait partie du tectum de la partie dorsale du diencéphale. Il est la cible de la plupart des relais auditifs de la voie ascendante ; ainsi il reçoit des projections du complexe olivaire supérieur, des noyaux cochléaires et des noyaux du lemnisque latéral, la plupart du temps ipsilatérales et contralatérales. Il est considéré comme un centre de traitement de l'information spatiale, destiné à fournir une carte spatiotopique de l'environnement auditif. Certains neurones montrent une sélectivité très fine en fréquence, tandis que d'autres sont insensibles aux sons stables et ne répondent qu'aux modulations d'amplitude pour certains ou de fréquence pour d'autres. Le colliculus inférieur poursuit donc l'analyse de la structure du signal sonore amorcée lors des voies auditives ascendantes. Il participe parallèlement au traitement de l'information spatiale et est plus sensible que le complexe olivaire aux disparités interaurales de phase et d'intensité. Certains neurones sont également sensibles au sens de déplacement d'une source sonore dans le plan médian, aussi peut-on penser que le colliculus inférieur intervient dans l'orientation de la tête ou des yeux en direction du son (Galifret, 1991).

Corps genouillé médian Le corps genouillé médian† (*medial geniculate body* ; MGB) est un relais thalamique recevant des projections du colliculus inférieur par deux voies séparées, l'une ventrale et l'autre dorsale. Ce relais présente une organisation laminaire en couches concentriques

parallèles à la surface, les couches superficielles correspondant aux basses fréquences et les couches profondes aux hautes fréquences. La sélectivité tonale reste bonne bien qu'inférieure à celle observée dans le colliculus inférieur. Certaines cellules présentent une sensibilité spécifique à des sons complexes particuliers. Beaucoup de cellules reçoivent des informations des deux oreilles et sont sensibles aux disparités d'intensité ou de phase, comme dans le colliculus inférieur. Les parties dorsale et médiale du MGB présentent des cellules moins sensibles à l'audition et effectuent le traitement d'informations en provenance d'autres modalités sensorielles ; ce sont des aires considérées comme multisensorielles. La partie dorsale projette sur le cortex auditif secondaire, tandis que la partie ventrale projette sur le cortex auditif primaire (Galifret, 1991).

1.3.2 Cortex auditif

Les projections du corps genouillé médian sur les aires corticales auditives sont le point terminal des voies auditives. Sur la base d'études cytoarchitectoniques des subdivisions du cortex auditif (Von Economo et Horn, 1930; Galaburda et Sanides, 1980; Morosan *et al.*, 2001; Rademacher *et al.*, 2001), trois aires corticales ont été définies : le cortex auditif primaire, le cortex auditif secondaire et les aires auditives associatives (ou cortex auditif tertiaire)

Cortex auditif primaire

Le cortex auditif primaire† (AI) est situé à l'intérieur du sillon latéral, le long de la face ventrale du lobe temporal supérieur, dans le gyrus temporal transverse† (gyrus de Heschl†) ; il correspond à l'aire 41 de Brodmann. Le cortex auditif primaire contient des neurones qui répondent à une gamme de fréquences comprises entre 20 et 20000 Hz et permettent l'identification de différentes caractéristiques des sons comme l'intensité, la hauteur ou le timbre. La réponse des neurones du cortex auditif primaire à des sons sinusoïdaux de fréquences spécifiques est très sélective et binaurale. Certaines cellules ont une réponse binaurale alors que d'autres sont excités par la stimulation d'une oreille et inhibées par la stimulation de l'oreille opposée. Le cortex auditif primaire est organisé en colonnes verticales organisées suivant une représentation tonotopique, chacune de ces colonnes étant sensible à une fréquence particulière.

Cortex auditif secondaire

Le cortex auditif primaire envoie des projections vers le cortex auditif secondaire†, par lequel il est entouré. Il recouvre le bord supérieur de la surface latérale du lobe temporal et correspond aux aires 42 et 22 de Brodmann. Le planum temporale† (PT) est l'une des principales régions du cortex auditif secondaire et recouvre principalement l'aire 42 de Brodmann, une petite partie touchant l'aire 22 de Brodamann. Il est situé à l'intérieur du sillon latéral, sur la face supérieure du lobe temporal. Le PT était au début considéré comme une aire intervenant essentiellement dans le traitement du langage, mais de récentes études fonctionnelles montrent qu'il joue également un rôle dans le traitement des sons et de la musique.

Cortex auditif associatif

Le cortex auditif associatif† regroupe l'ensemble des aires auditives associatives, qui correspondent approximativement à l'aire 22 de Brodmann, sont situées sur le lobe temporal, sur le gyrus temporal supérieur (*superior temporal gyrus* ; STG) et le sillon temporal supérieur (*superior temporal sulcus* ; STS). Elles sont situées latéralement et ventralement par rapport au cortex auditif primaire. Le cortex auditif associatif du STG projette vers le cortex multimodal

du STS, qui reçoit lui-même des projections des aires somesthésiques et visuelles (Baylis *et al.*, 1987; Desimone et Ungerleider, 1986). De nombreuses études ont établi le rôle des aires auditives associatives dans le traitement des informations auditives (Cansino *et al.*, 1994; Binder *et al.*, 1994, 1996, 1997, 2000; Zatorre *et al.*, 1992; Démonet *et al.*, 1992; Jancke *et al.*, 2002).

On pourra se référer en particulier à *« Intégration neuronale de la parole et du langage »*, ch. 4, p. 79 pour une description des aires auditives associatives impliquées dans la perception de la parole et du langage. Les implications multisensorielles des aires auditives associatives seront quant à elles décrites dans *« Intégration multisensorielle »*, ch. 5, p. 109. Enfin, le rôle des régions auditives en cas de déprivation sensorielle et après implantation cochléaire sera détaillé dans *« Réorganisation neurofonctionnelle après une altération sensorielle »*, ch. 8, p. 155 et *« Réorganisation neurofonctionnelle après implantation cochléaire »*, ch. 9, p. 191.

1.3.3 Anatomie fonctionnelle du cortex auditif

Tonotopie

Le cortex auditif, comme les autres cortex sensoriels, est organisé en cartes représentationnelles topographiques. La principale dimension autour de laquelle est organisée le cortex auditif est celle de la fréquence sonore. Ainsi on retrouve une représentation tonotopique depuis la membrane basilaire de la cochlée (voir *« Fonctionnement de la cochlée »*, § 1.1.2, p. 22) jusqu'aux régions corticales auditives. On pourra par ailleurs consulter *« Neuroplasticité intramodale »*, sect. 8.2, p. 161 pour une revue des effets d'une déprivation, d'une altération ou d'un entraînement perceptif sur l'organisation tonotopique.

Une description très détaillée des représentations tonotopiques corticales a été obtenue chez les primates non-humains grâce à des enregistrements électrophysiologiques. Le modèle couramment accepté distingue (Schonwiesner *et al.*, 2002) : 1° une région centrale, le cœur auditif (*core*) ; 2° un ensemble de régions directement adjacentes au cœur, la ceinture auditive (*auditory belt*) ; 3° un ensemble de régions adjacentes et latérales à la ceinture auditive, la para-ceinture auditive (*auditory parabelt*). Le *cœur* est caractérisé par une couche IV granulaire dense et comprend plusieurs cartes tonotopiques avec des gradients de fréquences (Merzenich et Brugge, 1973; Imig *et al.*, 1977; Morel et Kaas, 1992; Morel *et al.*, 1993; Hackett *et al.*, 1998; Kaas et Hackett, 1998, 2000). Ces cartes sont constituées de bandes isofréquentielles orientées médio-latéralement, alignées sur la face inférieure du sillon latéral depuis les hautes fréquences postéro-médiales jusqu'aux basses fréquences antéro-latérales. La *ceinture auditive* entoure le cœur et est composée de plus de sept aires corticales cytoarchitectoniquement distinctes, dont certaines semblent être également organisées de façon tonotopique (Pandya et Sanides, 1973; Kaas et Hackett, 1998, 2000). Ces régions, qui permettent un second niveau de traitement auditif et répondent à des sons plus complexes que pour le cœur, reçoivent peu d'afférences thalamiques et dépendent en grande partie des afférences en provenance du cœur auditif (Rauschecker, 1997). Enfin, la *para-ceinture auditive*, qui permet un troisième niveau de traitement auditif, est située latéralement à la ceinture auditive, sur la surface dorsale et dorso-latérale du gyrus temporal supérieur (Pandya et Sanides, 1973; Rauschecker *et al.*, 1995; Rauschecker, 1997).

Dans le cortex auditif humain, le principe de la représentation tonotopique des sons est confirmé par plusieurs études en neuroimagerie fonctionnelle en MEG (Romani *et al.*, 1982; Pantev *et al.*, 1988, 1989b, 1995), en TEP (Lauter *et al.*, 1985; Lockwood *et al.*, 1999) et en IRMf (Wessinger *et al.*, 1997; Strainer *et al.*, 1997; Bilecen *et al.*, 1998; Talavage *et al.*, 2000). Dans l'ensemble, ces études indiquent une représentation des hautes fréquences dans les parties les plus postéro-médiales, les basses fréquences étant représentées dans les parties les plus antéro-latérales. Une autre étude, réalisée par implantation de micro-électrodes dans le gyrus de Heschl droit d'un

patient épileptique, a permis d'évaluer directement les réponses fréquentielles du cortex auditif et d'observer de meilleures réponses aux hautes fréquences dans les parties postéro-médiales du cortex auditif primaire, tandis que les meilleurs réponses aux basses fréquences étaient situées dans les parties antéro-latérales (Howard *et al.*, 1996).

Toutefois il existe une importante variabilité interindividuelle du gyrus de Heschl, ce qui rend difficile la comparaison entre les cartes tonotopiques individuelles (Schonwiesner *et al.*, 2002). Les progrès des connaissances dans le domaine tendent à privilégier l'hypothèse d'une diversité des cartes tonotopiques dans le cortex auditif humain, chacune d'entre elles jouant un rôle spécifique dans la décomposition spectrale de la parole et d'autres sons complexes (Talavage *et al.*, 2000).

Spatiotopie et amplitopie

On dispose de preuves concernant l'existence d'un substrat neuronal de la spatiotopie dans les relais auditifs corticaux et thalamiques. Au niveau thalamique, des noyaux cérébraux comme le colliculus supérieur jouent probablement un rôle clé dans la localisation spatiale auditive chez les amphibiens, les oiseaux, et jusqu'à un certain point chez les mammifères (Rauschecker, 2004). La présence de cartes topgraphiques de l'espace auditif dans le colliculus supérieur a en effet été démontrée dans le cas des chouettes (Knudsen et Konishi, 1978) ainsi que des furets et des cobayes (Palmer et King, 1982).

Concernant les centres corticaux, des lésions du cortex auditif dans plusieurs espèces de mammifères, incluant les humains, révèlent d'importants déficits comportementaux dans la localisation sonore (Jenkins et Merzenich, 1984; Heffner et Heffner, 1990; Beitel et Kaas, 1993). Alors que d'autres études supposaient que le cortex auditif primaire était le centre nerveux de traitement des informations auditives spatiales, des études plus récentes montrent une probable implication des aires auditives associatives (Korte et Rauschecker, 1993; Middlebrooks *et al.*, 1994). L'analyse de la spécialisation fonctionnelle des aires corticales auditives chez le singe rhésus suggère l'existence d'une chaîne de traitement de la spatialité auditive qui débuterait au niveau de A1 et projetterait ensuite postérieurement et dorsalement jusqu'aux aires postérieures pariétales et préfrontales dorso-latérales (Rauschecker, 1998a,b; Romanski *et al.*, 1999; Rauschecker et Tian, 2000; Recanzone, 2000; Tian *et al.*, 2001). Ces aires, initialement conçues comme des aires dédiées au traitement de la spatialité spécifiquement visuelle, seraient donc également impliquées dans le traitement de la spatialité auditive (Rauschecker, 1998a,b).

Ces résultats provenant des études animales sont maintenant confortés par un nombre croissant d'études en neuroimagerie fonctionnelle chez l'humain, qui montrent que les sons provenant de l'hémichamp contralatéral sont plus excitants dans un hémisphère donné et confirment l'intervention des régions pariétales et auditives postérieures dans la localisation spatiale (Griffiths *et al.*, 1998; Bushara *et al.*, 1999; Weeks *et al.*, 1999; Alain *et al.*, 2001; Maeder *et al.*, 2001; Warren *et al.*, 2002; Zatorre *et al.*, 2002; Hall, 2003).

Par ailleurs, quelques études ont recherché une possible organisation amplitopique dans les régions auditives. Une telle organisation correspondrait à des cartes sensorielles neuronales distribuées spatialement en fonction de l'intensité sonore, avec des neurones répondant sélectivement à une intensité donnée. Par potentiels évoqués magnétiques, Pantev *et al.* (1989a) ont suggéré l'existence d'une telle représentation amplitopique dans le cortex auditif. Ces résultats semblent corroborés par des études ultérieures en IRMf, qui ont révélé une activation corticale en réseau bidimensionnel lors de l'augmentation de l'intensité sonore (Bilecen *et al.*, 2002). Ces activations ont été observées à l'intérieur du gyrus temporal transverse (gyrus de Heschl), avec une réponse pour des niveaux sonores croissants s'étendant du bord ventral au bord dorsal et de la partie

latérale à la partie médiale. Ce maillage latéro-médial serait donc une carte topographique des intensités sonores, analogue à l'organisation tonotopique du cortex auditif.

Mémoire auditive à court terme

Une propriété importante du système auditif humain est la mémoire auditive[†], qui implique les aptitudes à percevoir et traiter les informations acoustiques, dont notamment la voix, les sons environnementaux, la musique et d'autres sons complexes. Ces informations doivent être stockées en mémoire et leur rappel doit pouvoir intervenir de façon aussi bien volontaire qu'automatique. La mémoire auditive met en jeu de nombreuses tâches, incluant l'anticipation, l'écoute, le traitement, la mémorisation et le rappel, souvent toutes effectuées de façon simultanée. L'essentiel des recherches concernant la mémoire auditive ont été menées dans le domaine de la psychophysique. Les théories psychophysiques établissent ainsi une distinction entre la mémoire auditive sensorielle (mémoire échoïque) et la mémoire auditive à court terme.

Par définition, la mémoire auditive échoïque est un système qui capte des stimuli auditifs et les maintient pour une courte période de temps dans la mémoire sensorielle, et représente donc les niveaux de traitement les plus bas en matière de mémoire sensorielle. De nombreuses études comportementales et psychophysiques ont établi les propriétés spécifiques de la mémoire échoïque (Neisser, 1967; Darwin *et al.*, 1972; Cowan, 1984). La mémoire auditive sensorielle semble permettre à un stimulus de rester en mémoire en vue d'un traitement supplémentaire, qui pourrait être nécessaire pour une intégration des informations auditives en fonction du temps et de la position et pour une sélection des stimuli prégnants. Les études sur la mémoire échoïque suggèrent une similarité entre celle-ci et sa contrepartie visuelle, la mémoire iconique. Bien que la mémoire échoïque dure approximativement 4 secondes tandis que la mémoire iconique n'en dure qu'une seule, la mémoire échoïque ne semble pas conserver autant d'information que la mémoire iconique (Gaab, 2004).

Une méthode couramment utilisée pour étudier la mémoire auditive sensorielle est le paradigme du *mismatch negativity* (MMN), que l'on pourrait traduire approximativement par « négativité incongruente ». Durant les années 1970, de nombreuses études ont montré que le MMN apparaissait à la suite de la présentation d'un son incongruent dans une séquence cohérente de stimuli auditifs. Näätänen *et al.* (1978) suggèrent que le MMN est provoqué par une incongruence neuronale entre le stimulus incongruent et une trace sensorielle en mémoire, associée aux stimuli auditifs précédents. Comme le MMN peut également être remarqué dans un état d'inattention, il semblerait que les mécanismes de la mémoire échoïque opèrent sans traitement cognitif spécifique de l'attention. Comme les générateurs de MMN peuvent être localisés dans le cerveau, les corrélats neuronaux et les aires corticales associées à la mémoire auditive sensorielle peuvent être identifiés.

Durant les dix dernières années, le MMN a été appliqué dans de nombreux domaines de recherche, comme la perception des sons complexes (Schroger *et al.*, 1994; Alho *et al.*, 1996; Ioannides *et al.*, 2003), la perception de la parole et du langage chez les adultes et les nouveaux-nés (Winkler *et al.*, 1999; Cheour *et al.*, 2002), la perception de la musique (Tervaniemi *et al.*, 1997) et les déficits de mémoire sensorielle auditive dans des populations cliniques diversifiées (Näätänen et Escera, 2000; Näätänen, 2003). On s'accorde aujourd'hui sur l'importance essentielle du MMN dans le compréhension des mécanismes liés à la mémoire échoïque, même si le MMN n'est pas nécessairement le corrélat physiologique direct de la mémoire échoïque (Näätänen *et al.*, 1989b,a; Jaaskelainen *et al.*, 1999).

Bibliographie

Alain, C., Arnott, S. R., Hevenor, S., Graham, S. et Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proc. Natl. Acad. Sci. U.S.A.*, 98(21):12301–12306.

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., Knuutila, J. et Näätänen, R. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology*, 33(4):369–375.

Baylis, G. C., Rolls, E. T. et Leonard, C. M. (1987). Functional subdivisions of the temporal lobe neocortex. *J. Neurosci.*, 7(2):330–342.

Bear, M. F., Connors, B. W. et Paradiso, M. A. (1997). *In* Pradel, Masson-Williams et Wilkins, éditeurs : *Neurosciences : à la découverte du cerveau*. Paris, France.

Beitel, R. E. et Kaas, J. H. (1993). Effects of bilateral and unilateral ablation of auditory cortex in cats on the unconditioned head orienting response to acoustic stimuli. *J. Neurophysiol.*, 70(1):351–369.

Bilecen, D., Scheffler, K., Schmid, N., Tschopp, K. et Seelig, J. (1998). Tonotopic organization of the human auditory cortex as detected by BOLD-FMRI. *Hear. Res.*, 126(1-2):19–27.

Bilecen, D., Seifritz, E., Scheffler, K., Henning, J. et Schulte, A. C. (2002). Amplitopicity of the human auditory cortex : an fMRI study. *Neuroimage*, 17(2):710–718.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N. et Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex*, 10(5):512–528.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M. et Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci.*, 17(1):353–362.

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M. et Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, 119(Pt 4):1239–1247.

Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., Estkowski, L. D., Goldstein, M. D., Haughton, V. M. et Hyde, J. S. (1994). Functional magnetic resonance imaging of human auditory cortex. *Ann. Neurol.*, 35(6):662–672.

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P. et Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.*, 2(8):759–766.

Cansino, S., Williamson, S. J. et Karron, D. (1994). Tonotopic organization of human auditory association cortex. *Brain Res.*, 663(1):38–50.

Cheour, M., Shestakova, A., Alku, P., Ceponiene, R. et Näätänen, R. (2002). Mismatch negativity shows that 3-6-year-old children can learn to discriminate non-native speech sounds within two months. *Neurosci. Lett.*, 325(3):187–190.

Cowan, N. (1984). On short and long auditory stores. *Psychol. Bull.*, 96(2):341–370.

Darwin, C. J., Turvey, M. T. et Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure. Evidence for brief auditory store. *Cognit. Psychol.*, 3:255–267.

Desimone, R. et Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J. Comp. Neurol.*, 248(2):164–189.

Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A. et Frackowiak, R. S. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115:6.

Fletcher, H. (1940). Auditory patterns. *Rev. Mod. Phys.*, 12:47–65.

Gaab, N. (2004). *The auditory cortex : perception, memory, plasticity and the influence of musicianship.* Thèse de doctorat, Université de Zürich.

Galaburda, A. M. et Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.*, 190:597–610.

Galifret, Y. (1991). Acoustique physiologique. *In Encyclopaedia universalis*, volume 1, pages 198–210. Encyclopaedia universalis.

Griffiths, T. D., Rees, G., Rees, A., Green, G. G., Witton, C., Rowe, D., Buchel, C., Turner, R. et Frackowiak, R. S. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nat. Neurosci.*, 1(1):74–79.

Grothe, B. et Park, T. J. (2000). Structure and function of the bat superior olivary complex. *Microsc. Res. Tech.*, 51(4):382–402.

Hackett, T. A., Stepniewska, I. et Kaas, J. H. (1998). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J. Comp. Neurol.*, 394(4):475–495.

Hall, D. A. (2003). Auditory pathways : are "what" and "where" appropriate ? *Curr. Biol.*, 13(10):406–408.

Heffner, H. E. et Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in japanese macaques. *J. Neurophysiol.*, 64(3):915–931.

Howard, 3rd, M. A., Volkov, I. O., Abbas, P. J., Damasio, H., Ollendieck, M. C. et Granner, M. A. (1996). A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res.*, 724(2):260–264.

Imig, T. J., Ruggero, M. A., Kitzes, L. M., Javel, E. et Brugge, J. F. (1977). Organization of auditory cortex in the owl monkey (Aotus trivirgatus). *J. Comp. Neurol.*, 171(1):111–128.

Ioannides, A. A., Popescu, M., Otsuka, A., Bezerianos, A. et Liu, L. (2003). Magnetoencephalographic evidence of the interhemispheric asymmetry in echoic memory lifetime and its dependence on handedness and gender. *Neuroimage*, 19(3):1061–1075.

Jaaskelainen, I. P., Hautamaki, M., Näätänen, R. et Ilmoniemi, R. J. (1999). Temporal span of human echoic memory and mismatch negativity : revisited. *Neuroreport*, 10(6):1305–1308.

Jancke, L., Wustenberg, T., Scheich, H. et Heinze, H. J. (2002). Phonetic perception and the temporal cortex. *Neuroimage*, 15(4):733–746.

Jenkins, W. M. et Merzenich, M. M. (1984). Role of cat primary auditory cortex for sound-localization behavior. *J. Neurophysiol.*, 52(5):819–847.

Kaas, J. H. et Hackett, T. A. (1998). Subdivisions of auditory cortex and levels of processing in primates. *Audiol. Neurootol.*, 3(2-3):73–85.

Kaas, J. H. et Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.*, 97(22):11793–11799.

Knudsen, E. I. et Konishi, M. (1978). A neural map of auditory space in the owl. *Science*, 200(4343):795–797.

Korte, M. et Rauschecker, J. P. (1993). Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *J. Neurophysiol.*, 70(4):1717–1721.

Lauter, J. L., Herscovitch, P., Formby, C. et Raichle, M. E. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear. Res.*, 20(3):199–205.

Lockwood, A. H., Salvi, R. J., Coad, M. L., Arnold, S. A., Wack, D. S., Murphy, B. W. et Burkard, R. F. (1999). The functional anatomy of the normal human auditory system : responses to 0.5 and 4.0 kHz tones at varied intensities. *Cereb. Cortex*, 9(1):65–76.

Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., Pittet, A. et Clarke, S. (2001). Distinct pathways involved in sound recognition and localization : a human fMRI study. *Neuroimage*, 14(4):802–816.

Merzenich, M. M. et Brugge, J. F. (1973). Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res.*, 50(2):275–296.

Middlebrooks, J. C., Clock, A. E., Xu, L. et Green, D. M. (1994). A panoramic code for sound location by cortical neurons. *Science*, 264(5160):842–844.

Moore, B. C. J. (1989a). Frequency selectivity, masking and the critical band. *In An introduction to the psychology of hearing*, pages 84–136. Academic Press Limited, London, 3^{e} édition.

Moore, B. C. J. (1989b). The nature of the sound and the structure of the auditory system. *In An introduction to the psychology of hearing*, pages 1–46. Academic Press Limited, London, 3^{e} édition.

Moore, B. C. J. (1998). *Cochlear hearing loss*. Whurr, London.

Morel, A., Garraghty, P. E. et Kaas, J. H. (1993). Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J. Comp. Neurol.*, 335(3):437–459.

Morel, A. et Kaas, J. H. (1992). Subdivisions and connections of auditory cortex in owl monkeys. *J. Comp. Neurol.*, 318(1):27–63.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T. et Zilles, K. (2001). Human primary auditory cortex : cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage*, 13:684–701.

Neisser, U. (1967). *Cognitive Psychology*. New York, NY : Apelton-Century-Crofts.

Näätänen, R. (2003). Mismatch negativity : clinical research and possible applications. *Int. J. Psychophysiol.*, 48(2):179–188.

Näätänen, R. et Escera, C. (2000). Mismatch negativity : clinical and other applications. *Audiol. Neurootol.*, 5(3-4):105–110.

Näätänen, R., Gaillard, A. W. et Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.*, 42(4):313–329.

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K. et Sams, M. (1989a). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain ? *Neurosci. Lett.*, 98(2):217–221.

Näätänen, R., Paavilainen, P. et Reinikainen, K. (1989b). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man ? *Neurosci. Lett.*, 107(1-3):347–352.

Palmer, A. R. et King, A. J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature*, 299(5880):248–249.

Pandya, D. N. et Sanides, F. (1973). Architectonic parcellation of the temporal operculum in rhesus monkey and its projection pattern. *Z. Anat. Entwicklungsgesch.*, 139(2):127–161.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G. et Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalogr. Clin. Neurophysiol.*, 94(1):26–40.

Pantev, C., Hoke, M., Lehnertz, K. et Lutkenhoner, B. (1989a). Neuromagnetic evidence of an amplitopic organization of the human auditory cortex. *Electroencephalogr. Clin. Neurophysiol.*, 72(3):225–231.

Pantev, C., Hoke, M., Lehnertz, K., Lutkenhoner, B., Anogianakis, G. et Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalogr. Clin. Neurophysiol.*, 69(2):160–170.

Pantev, C., Hoke, M., Lutkenhoner, B. et Lehnertz, K. (1989b). Tonotopic organization of the auditory cortex : pitch versus frequency representation. *Science*, 246(4929):486–488.

Park, T. J., Grothe, B., Pollak, G. D., Schuller, G. et Koch, U. (1996). Neural delays shape selectivity to interaural intensity differences in the lateral superior olive. *J. Neurosci.*, 16(20):6554–6566.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *J. Acoust. Soc. Am.*, 59(3):640–654.

Pujol, R., Réclar-Enjalbert, V., Pujol, T., Blatrix, S., Minary, P., Roméro, C., Gressel, F., Puel, J. L., Lenoir, M., Lavigne-Rebillard, M., Eybalin, M. et Rebillard, G. (1999). Promenade autour de la cochlée. Site internet. http://www.iurc.montp.inserm.fr/cric/audition/start.htm.

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H. J. et Zilles, K. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage*, 13:669–683.

Rauschecker, J. P. (1997). Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol. Suppl.*, 532:34–38.

Rauschecker, J. P. (1998a). Cortical processing of complex sounds. *Curr. Opin. Neurobiol.*, 8(4):516–521.

Rauschecker, J. P. (1998b). Parallel processing in the auditory cortex of primates. *Audiol. Neurootol.*, 3(2-3):86–103.

Rauschecker, J. P. (2004). Cross-modal consequences of visual deprivation in animals. *In* Calvert, G. A., Spence, C. et Stein, B. E., éditeurs : *The Handbook of Multisensory Processes*, pages 695–702. Cambridge, MA : MIT Press.

Rauschecker, J. P. et Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.*, 97(22):11800–11806.

Rauschecker, J. P., Tian, B. et Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268(5207):111–114.

Recanzone, G. H. (2000). Spatial processing in the auditory cortex of the macaque monkey. *Proc. Natl. Acad. Sci. U.S.A.*, 97(22):11829–11835.

Romani, G. L., Williamson, S. J. et Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. *Science*, 216(4552):1339–1340.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S. et Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.*, 2(12):1131–1136.

Schonwiesner, M., von Cramon, D. Y. et Rubsamen, R. (2002). Is it tonotopy after all? *Neuroimage*, 17(3):1144–1161.

Schroger, E., Paavilainen, P. et Näätänen, R. (1994). Mismatch negativity to changes in a continuous tone with regularly varying frequencies. *Electroencephalogr. Clin. Neurophysiol.*, 92(2):140–147.

Strainer, J. C., Ulmer, J. L., Yetkin, F. Z., Haughton, V. M., Daniels, D. L. et Millen, S. J. (1997). Functional MR of the primary auditory cortex : an analysis of pure tone activation and tone discrimination. *AJNR Am. J. Neuroradiol.*, 18(4):601–610.

Talavage, T. M., Ledden, P. J., Benson, R. R., Rosen, B. R. et Melcher, J. R. (2000). Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hear. Res.*, 150(1-2):225–244.

Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K. et Näätänen, R. (1997). The musical brain : brain waves reveal the neurophysiological basis of musicality in human subjects. *Neurosci. Lett.*, 226(1):1–4.

Tian, B., Reser, D., Durham, A., Kustov, A. et Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science*, 292(5515):290–293.

Tollin, D. J. (2003). The lateral superior olive : a functional role in sound source localization. *Neuroscientist*, 9(2):127–143.

Von Economo, C. et Horn, L. (1930). Über Windungsrelief MaSSe und Rindenarchitektonik der Supratemporalfläche, ihre individuellen und ihre Seitenunterschiede. *Z. Neurol. Psychiatrie*, 130:678–757.

Warren, J. D., Zielinski, B. A., Green, G. G., Rauschecker, J. P. et Griffiths, T. D. (2002). Perception of sound-source motion by the human brain. *Neuron*, 34(1):139–148.

Weeks, R. A., Aziz-Sultan, A., Bushara, K. O., Tian, B., Wessinger, C. M., Dang, N., Rauschecker, J. P. et Hallett, M. (1999). A PET study of human auditory spatial processing. *Neurosci. Lett.*, 262(3):155–158.

Wessinger, C. M., Bounocore, M. H., Kussmaul, C. L. et Mangun, G. R. (1997). Tonotopy in human auditory cortex examined with functional magnetic resonance imaging. *Hum. Brain Mapp.*, 5:18.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R. J. et Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 36(5):638–642.

Yin, T. C. et Chan, J. C. (1990). Interaural time sensitivity in medial superior olive of cat. *J. Neurophysiol.*, 64(2):465–488.

Zatorre, R. J., Bouffard, M., Ahad, P. et Belin, P. (2002). Where is "where" in the human auditory cortex ? *Nat. Neurosci.*, 5(9):905–909.

Zatorre, R. J., Evans, A. C., Meyer, E. et Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058):846–849.

Chapitre 2

Production de la parole

Sommaire

2.1 Appareil phonatoire 42
2.1.1 Soufflerie pulmonaire 42
2.1.2 Le larynx 43
2.1.3 L'espace supralaryngé 45
2.2 Aspects articulatoires de la phonation 46
2.2.1 Généralités sur l'articulation 46
2.2.2 Notation phonétique 48
2.2.3 Caractéristiques articulatoires des voyelles 48
2.2.4 Caractéristiques articulatoires des consonnes 50
2.2.5 Les glides 51
2.3 Caractéristiques acoustiques des phonèmes 52
2.3.1 Structure acoustique des phonèmes 52
2.3.2 Caractéristiques acoustiques des voyelles 54
2.3.3 Caractéristiques acoustiques des consonnes 54
Bibliographie 57

Introduction

La compréhension des processus physiologiques qui permettent à l'être humain de bien percevoir les paroles émises par ses semblables joue un rôle primordial dans l'étude des patients sourds implantés cochléaires. En effet, les implants cochléaires (et de façon plus générale les prothèses auditives) sont essentiellement destinés à rétablir l'aptitude des personnes sourdes à utiliser correctement la communication parlée. Il nous faut donc aborder le thème de la parole, ce signal physique à la fois acoustique et optique, véhiculant aussi bien les sons de la voix que les mouvements des lèvres du locuteur.

Dans la suite des évènements qui permettent à une personne de percevoir la parole, les sons du langage sont initialement générés par un locuteur. Ils sont éventuellement captés par un microphone puis enregistrés ou transformés, pour être le cas échéant réémis par des dispositifs technologiques (téléphone, télévision, radio, magnétophone, etc.). Ils

se propagent ensuite sous forme d'ondes acoustiques dans l'air puis parviennent au système auditif, qui transforme ces signaux acoustiques en signaux électriques communiqués aux centres nerveux de l'audition. La personne en question prend alors connaissance du contenu acoustique de la parole. Ce contenu purement physique, auquel n'a pas encore été associée une signification particulière, est ensuite traité et interprété par les centres nerveux du langage, ce qui permet de lui donner un contenu sémantique et rend possible une interprétation de ce qui a été perçu.

La plupart des situations de communication quotidienne (notamment conversation en face-à-face) mettent également en jeu la modalité visuelle. Les signaux acoustiques contenant les sons du langage sont dans ce cas systématiquement couplés avec des signaux optiques renseignant sur les mouvements articulatoires de la langue et des lèvres, ainsi que les mouvements et expressions faciales et corporelles. De manière similaire, ces signaux sont captés par le système visuel puis transformés en signaux électriques communiqués aux centres nerveux de la vision. Il est très important de prendre entièrement en compte ces aspects visuels de la parole, comme nous allons le voir dans la suite.

On peut donc distinguer deux étapes dans l'acquisition du contenu phonétique de la parole : sa production par le locuteur au niveau de l'appareil phonatoire, et plus particulièrement du larynx, puis sa perception par l'interlocuteur au niveau des systèmes auditif et visuel et des aires corticales associées. Le présent chapitre est consacré à la production des sons du langage, qui constitue l'étape initiale de la communication parlée. Nous décrivons ici l'anatomie de l'appareil phonatoire humain et l'éventail des mouvements articulatoires possibles, ce qui permet d'expliquer la répartition des phonèmes en différentes catégories acoustiques et articulatoires clairement définies.

2.1 Appareil phonatoire[1]

La production articulatoire des sons du langage chez l'être humain est un processus complexe qui met en œuvre des mécanismes neurologiques et physiologiques. Un nombre important d'organes et de muscles interviennent dans l'émission de la parole. On décompose usuellement le fonctionnement de l'appareil phonatoire humain en trois étapes successives : la production d'une colonne d'air au niveau des poumons, sa mise en vibration au niveau du larynx et des cordes vocales, ce qui a pour effet de générer un son, puis la modulation de ce son par les différents résonateurs acoustiques formés par les cavités supralaryngées.

2.1.1 Soufflerie pulmonaire

Les organes respiratoires sont constitués des poumons, du diaphragme et de la trachée. Le processus de la respiration normale comprend deux phases de durées sensiblement égales : l'inspiration et l'expiration. Lors de l'inspiration, le diaphragme s'abaisse et les muscles intercostaux se contractent en provoquant une expansion de la cage thoracique, ce qui a pour effet de créer un appel d'air qui permet d'emplir les poumons. Lors de l'expiration, le diaphragme se relâche et la cage thoracique se rétracte, provoquant une compression des poumons et l'expulsion des gaz respiratoires. Seule l'inspiration nécessite un travail moteur des muscles respiratoires, l'expiration étant due à la retombée passive de la cage thoracique.

La phonation humaine intervient majoritairement durant la phase expiratoire, qui fournit le flux d'air nécessaire à la production des sons laryngés. En particulier, le français et l'anglais

[1] d'après Abitbol (1991); Chiss *et al.* (1996)

utilisent exclusivement l'expiration. En conséquence, lors de la production de la parole, on peut observer une modification du cycle respiratoire. Alors que dans des conditions normales, l'inspiration et l'expiration se répartissent également le temps de respiration, on constate que lors de la phonation, seuls 15% du cycle respiratoire sont consacrés à l'inspiration. Par ailleurs, dans ces mêmes conditions normales, la pression expiratoire ne dépasse la pression atmosphérique que de 0,25%, tandis que cette surpression atteint 1% lors de la phonation.

Ces modifications du cycle respiratoire assurent une alimentation suffisante en air de l'appareil phonatoire pour permettre une génération efficace des sons de la parole. La maîtrise du débit et de la pression de l'air expiré lors de la phonation est relayée par les muscles de la sangle abdominale, qui permettent un contrôle précis du volume pulmonaire.

2.1.2 Le larynx

Le larynx[†] est un organe formé d'un assemblage de cartilages situés dans la partie moyenne du cou, juste au-dessus de la trachée, au-dessous de l'os hyoïde et de la base de la langue. Du point de vue fonctionnel, le larynx intervient lors de la déglutition, de la phonation et de certains modes particuliers de respiration, comme pendant une activité physique intense. On désigne par l'adjectif supralaryngé toute partie des voies respiratoires supérieures située au-dessus du larynx.

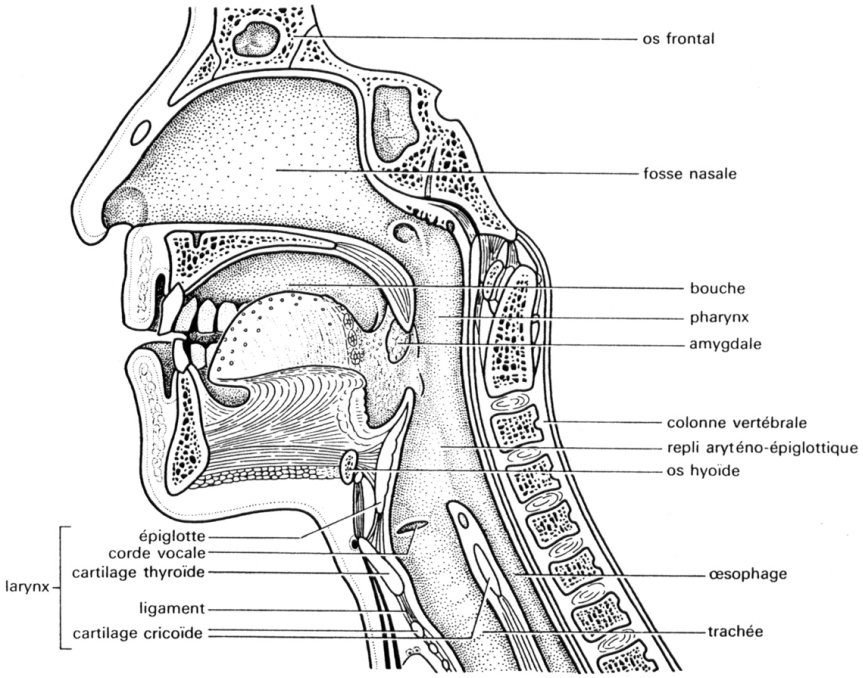


FIG. 2.1 – Organes de la parole vue en coupe sagittale. D'après Abitbol (1991).

La base du larynx est formée du cricoïde, un anneau cartilagineux horizontal différencié de la trachée, dont il surmonte le sommet. Le bord postérieur de cet anneau s'amincit en une petite saillie orientée verticalement. Au-dessus de cette saillie postérieure sont disposées symétriquement deux petites pyramides cartilagineuses, les aryténoïdes, dont les apophyses servent de point d'attache postérieur aux cordes vocales. Un système complexe de muscles permet aux aryténoïdes de glisser et de pivoter sur le cricoïde.

Sur le bord antérieur du cricoïde vient se baser le cartilage thyroïde, un volumineux cartilage bombé en forme de bouclier, dont la saillie antérieure est visible à travers la peau du cou et est communément appelée pomme d'Adam. Il est susceptible de pivoter légérement par rapport au cartilage cricoïde. Ce cartilage est muni à ses quatre coins d'apophyses verticales appelées respectivement cornes inférieures et cornes supérieures. La face postérieure du cartilage thyroïde sert de point d'attache antérieur aux cordes vocales ; le thyroïde permet également d'assurer leur protection.

L'épiglotte est un clapet cartilagineux dont la pointe inférieure est fixée au prolongement supérieur du cartilage thyroïde. Elle a pour fonction de couper la communication entre la trachée et les voies respiratoires supérieures, ce qui permet d'empêcher l'entrée de corps étrangers dans les poumons lors de la déglutition ou d'emprisonner l'air dans les poumons lors d'efforts physiques particuliers.

La paroi interne du larynx est tapissée d'une couche de tissu fibroélastique, le cône élastique, qui s'étend de l'intérieur des anneaux de la trachée artère jusqu'au-dessus du cartilage cricoïde, où il est successivement rétréci par deux doubles replis dirigés vers l'intérieur. Les replis inférieurs sont les cordes vocales[†], ou vraies cordes vocales. Celles-ci sont renforcées par des ligaments et par les muscles thyro-aryténoïdiens, également appelés muscles vocaux, qui assurent la tension longitudinale des cordes vocales. La glotte désigne l'espace libre délimité par les cordes vocales, la paroi interne du cartilage thyroïde et les apophyses antérieures des aryténoïdes. Les replis supérieurs sont appelés fausses cordes vocales, et les cavités délimitées de chaque côté du larynx par ces deux systèmes de replis sont dénommées ventricules de Morgagni.

Les muscles crico-thyroïdiens rattachent le cartilage thyroïde au cartilage cricoïde. Ce sont des tenseurs des cordes vocales. En se contractant, ils permettent au cricoïde de basculer légèrement vers l'arrière du larynx, entraînant avec lui les arythénoïdes et les cordes vocales, ce qui assure la tension longitudinale de ces dernières. Les muscles crico-aryténoïdiens relient les deux aryténoïdes au cartilage cricoïde. Les crico-aryténoïdiens postérieurs sont responsables de la constriction de la glotte, tandis que les crico-aryténoïdiens latéraux assurent sa dilatation. Enfin, les muscles inter-aryténoïdiens relient les deux cartilages aryténoïdes l'un à l'autre et contribuent aux mouvements de constriction de la glotte.

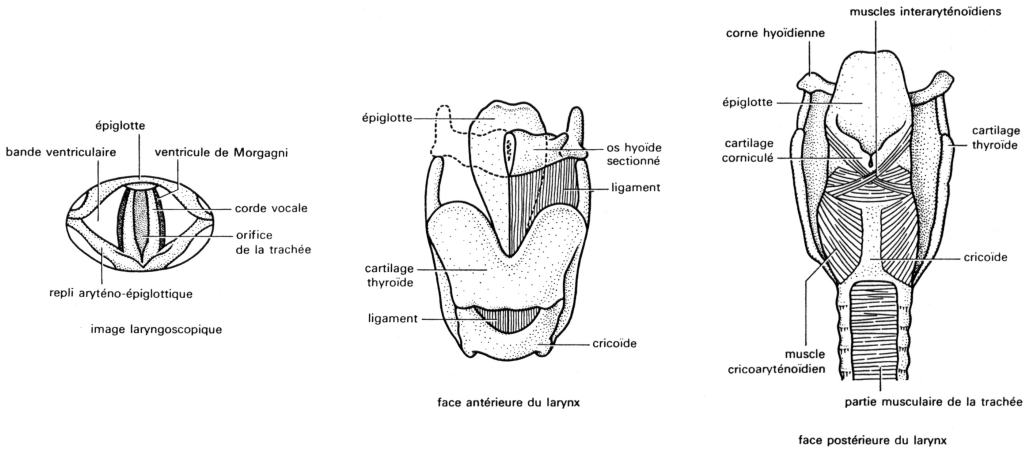


FIG. 2.2 – Structure anatomique du larynx. D'après Abitbol (1991).

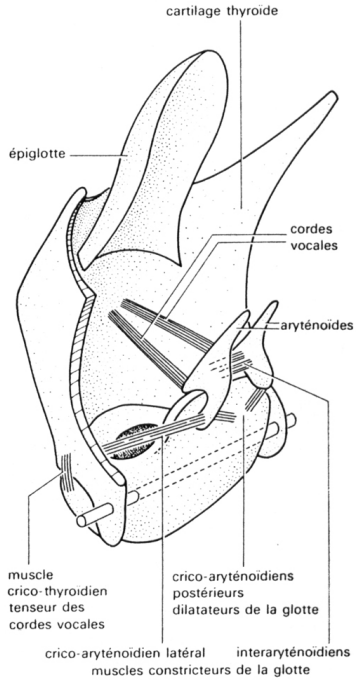


FIG. 2.3 – Schéma de la mécanique laryngée. D'après Abitbol (1991).

Génération du son dans le larynx

L'émission des sons laryngés se produit lorsque les cordes vocales entrent en vibration. La théorie myoélastique interprète cette entrée en vibration comme un phénomène essentiellement passif, dû à l'action mécanique du flux d'air en provenance des poumons. L'action des muscles tenseurs et adducteurs des cordes vocales provoque leur rapprochement et l'occlusion de la glotte. Lors de l'expiration, la pression subglottique des gaz respiratoires devient suffisamment élevée pour provoquer l'écartement des cordes vocales. Une certaine quantité d'air est alors libérée dans le pharynx, ce qui induit une dépression au niveau de la glotte et un abaissement de la pression subglottique. La résistance opposée par la colonne d'air n'étant plus suffisante, les cordes vocales reviennent à leur position occlusive initiale par simple élasticité. Le phénomène se reproduit alors de façon cyclique.

On obtient ainsi une alternance cyclique d'écartements et de rapprochements des cordes vocales. Ces battements périodiques engendrent une onde acoustique qui va se propager dans l'espace supralaryngé et faire entrer en résonance les différentes cavités délimitées par le conduit vocal, ce qui permet la production des sons de la parole.

Une autre théorie de production des sons laryngés, appelée théorie neuro-chronaxique et développée par Raoul Husson à partir de 1950, fait jouer au système nerveux un rôle prépondérant en expliquant l'écartement des cordes vocales par une commande nerveuse volontaire des muscles du larynx. Cette théorie a été progressivement abandonnée, s'accordant mal aux observations anatomiques et physiologiques.

2.1.3 L'espace supralaryngé

L'onde acoustique générée dans le larynx par l'entrée en vibration des cordes vocales se propage dans les ventricules de Morgagni puis dans les différentes cavités de l'espace supralaryngé†.

On distingue notamment la cavité pharyngale, la cavité buccale, les cavités nasales et éventuellement la cavité labiale.

Le pharynx† est un conduit musculo-membraneux situé au-dessus du larynx, en avant de la colonne vertébrale et en arrière de la langue, du voile du palais, de la luette et des fosses nasales. Il met en communication la bouche et les fosses nasales avec l'œsophage et le larynx. Il est susceptible de se contracter sous l'action des muscles qui le délimitent et peut être déformé par les mouvements de la base de la langue, du voile du palais, et de l'épiglotte. On appelle cavité pharyngale† le volume d'air délimité par le pharynx, susceptible de jouer un rôle de résonateur acoustique lors de la phonation.

La cavité buccale† est le volume d'air contenu dans la bouche, délimité antérieurement par les lèvres et les dents, postérieurement par la luette et le pharynx, inférieurement par la langue et supérieurement par le palais dur et le voile du palais. La langue†, constituée d'une quinzaine de muscles et dont la partie arrière est fixée à l'os hyoïde, joue un rôle fondamental dans la phonation de par sa grande flexibilité. Celle-ci lui permet en effet de modifier considérablement le volume de la cavité buccale et même de la diviser en deux cavités, l'une antérieure, située en avant de la langue, et l'autre postérieure, située en arrière de la langue.

Le voile du palais†, appelé également palais mou ou velum†, est la partie postérieure de la voûte de la cavité buccale. Il assure la séparation de la bouche et des fosses nasales en se relevant lors de la déglutiton. Il peut s'abaisser pendant la phonation et mettre en communication la cavité buccale et les fosses nasales. Les fosses nasales†, ou cavités nasales†, sont séparées l'une de l'autre par une cloison médiane plate et possèdent des volumes sensiblement constants.

Les sont susceptibles de resserrer ou de refermer l'orifice du conduit buccal. Lorsqu'elles sont arrondies ou projetées vers l'extérieur, elles délimitent avec les dents un petit volume appelé cavité labiale†.

2.2 Aspects articulatoires de la phonation[2]

L'articulation consiste pour le locuteur à déformer ses cavités supralaryngées par des mouvements musculaires adaptés, ce qui lui permet de modifier la structure spectrale et temporelle de l'onde acoustique émise par les cordes vocales et d'émettre ainsi des sons de parole. La phonétique articulatoire a pour but de décrire le plus précisément possible les stratégies de réalisation orale utilisées par les locuteurs.

2.2.1 Généralités sur l'articulation

Un phonème† est considéré comme la plus petite unité sonore pertinente dans une langue donnée. La production des phonèmes est liée à différents facteurs anatomiques. Le mode articulatoire† a trait à la qualité du passage de l'air dans l'espace supralaryngé, au-delà de la glotte. La réalisation des voyelles implique un passage libre de l'air dans cet espace, tandis que la réalisation des consonnes nécessite une gêne ou un arrêt du passage de l'air. Le passage des consonnes aux voyelles ne se fait pas de manière abrupte, mais sur un continuum, ce qui permet de distinguer des articulations intermédiaires comme les glides.

La sonorité, déterminée par une résonance orale ou nasale, est fonction de l'ouverture ou de la fermeture de l'accès vers les fosses nasales. Lors de la production de voyelles ou de consonnes nasales, le voile du palais s'abaisse et se décolle de la paroi pharyngale, ce qui permet le passage de l'air à la fois par le canal buccal et par les fosses nasales, ce qui confère aux sons une coloration

[2] d'après Autesserre (1991); Chiss *et al.* (1996)

particulière. Les voyelles et les consonnes produites sont alors dites nasales. Lorsque le voile du palais est relevé et bien accolé à la paroi pharyngale, l'air ne passe que par la cavité buccale, donnant naissance aux sons vocaliques et consonantiques dits oraux.

Le mode d'intervention des cordes vocales détermine le caractère sourd ou sonore des différentes articulations. Lorsque les cordes vocales vibrent, les sons sont dits voisés† ou sonores par opposition aux sons non voisés† ou sourds. La réalisation des voyelles implique la mise en vibration des cordes vocales. Pour les consonnes, l'absence ou la présence de vibration des cordes vocales détermine leur caractère sourd ou sonore. En français, toutes les consonnes nasales sont voisées. La distinction entre consonnes sourdes et sonores ne vaut que pour les consonnes orales.

La réalisation de toute articulation implique un rapprochement plus ou moins grand ou un contact franc entre deux points anatomiques, l'articulateur† et le lieu d'articulation†. Le lieu d'articulation se situe nécessairement dans la partie supérieure du canal buccal dans une zone allant de la lèvre supérieure jusqu'à la paroi pharyngale. C'est le point duquel l'articulateur se rapprochera ou avec lequel il entrera en contact. Les points d'articulation possibles sont :

Organe	**Nomenclature phonétique**
lèvres	labiale
dents	dentale
alvéoles	alvéolaire postalvéolaire ou alvéopalatale
palais dur	prépalatale médiopalatale postpalatale
voile du palais	prévélaire postvélaire
luette	uvulaire
pharynx	pharyngale
glotte	glottale ou pharyngale

TAB. 2.1 – Nomenclature phonétique des articulations selon la partie anatomique utilisée

L'articulateur est quant à lui constitué par la région inférieure du canal buccal. Il peut s'agir de la lèvre inférieure ou des différentes parties de la langue, subdivisée en régions apicale, prédorsale, médiodorsale, postdorsale et radicale de la pointe (apex) à la racine (radix).

On distingue ainsi trois catégories articulatoires. Les voyelles† peuvent être décrites comme des sons qui nécessitent la mise en vibration des cordes vocales et un libre passage de l'air dans le canal buccal, permis par un éloignement suffisant des articulateurs. Les consonnes† sont au contraire des sons générés lorsque le passage de l'air est perturbé par un resserrement plus ou moins marqué du canal buccal, pouvant aller jusqu'à sa fermeture complète pour les consonnes occlusives. Enfin les glides†, qui peuvent également être désignées par les termes ambigüs semi-consonnes† ou semi-voyelles†, correspondent à un état d'articulation intermédiaire entre les voyelles, pour lesquelles l'air passe de façon totalement libre dans le canal buccal, et les consonnes, pour lesquelles le flux d'air est fortement perturbé par le rapprochement des articulateurs.

2.2.2 Notation phonétique

L'étude de la parole nécessite une transcription des sons du langage dans un système d'écriture adapté. Or la plupart des systèmes d'écriture des langues du monde comportent une certaine ambigüité quant à la prononciation qui devrait leur être associée. Ainsi en français, dans le mot *soir*, la succession écrite des voyelles *o* et *i* ne correspond pas à la succession orale d'un *o* et d'un *i* indépendants. De même, dans les mots *taxi*, *hexagone*, *soixante* et *sixième*, la présence de la même lettre *x* dans l'écriture de chacun de ces mots se traduit par quatre réalisations orales différentes.

Les symboles correspondant aux lettres de l'alphabet apparaissent donc comme insuffisants pour établir des relations non ambigües entre la réalisation orale et la transcription écrite. On voit apparaître la nécessité d'introduire un nouveau système d'écriture qui assurerait une correspondance univoque entre l'écrit et l'oral, un son étant associé de façon unique à un symbole graphique et réciproquement. C'est le rôle que joue l'API (alphabet phonétique international)†, créé en 1888 par l'Association phonétique internationale. Ce système phonétique permet de transcrire l'ensemble des sons utilisés par les différentes langues du monde sous forme écrite.

i	l*i*t	y	l*u*
e	d*é*	ø	d*eu*x
ɛ	d*ai*s	œ	p*eu*r
a	t*a*	ə	d*e*
ɑ	t*a*s	ɑ̃	d*en*t
ɔ	s*o*rt	ɔ̃	d*on*
o	s*o*t	ɛ̃	br*in*
u	l*ou*p	œ̃	br*un*

TAB. 2.2 – Symboles de l'API utilisés pour les voyelles en français

p	*p*ont	f	*f*er	j	c*i*el
b	*b*on	v	*v*erre	ɥ	l*u*i
m	*m*on			w	L*ou*is
		s	*s*aisir		
t	*t*emps	z	sai*s*ir		
d	*d*ent				
n	*n*eige	ʃ	*ch*amp		
		ʒ	*J*ean		
k	*c*lair				
g	*g*ant	ʀ	*r*ond		
ɲ	a*gn*eau	l	*l*ong		

TAB. 2.3 – Symboles de l'API utilisés pour les consonnes et les semi-consonnes en français

2.2.3 Caractéristiques articulatoires des voyelles

Les voyelles† peuvent être décrites comme des sons qui nécessitent la mise en vibration des cordes vocales et un libre passage de l'air dans le canal buccal, permis par un éloignement suffisant des articulateurs.

On peut décrire les voyelles suivant leur lieu d'articulation, qui désigne essentiellement l'endroit de la voûte palatine vers lequel la langue gonfle sa masse musculaire. Les voyelles palatales, ou voyelles antérieures, sont produites lorsque la partie antérieure de la langue (zones apicale et pré-dorsale) s'élève vers le palais dur. Les voyelles vélaires, ou voyelles postérieures, sont à l'inverse produites lorsque la partie postérieure de la langue (zones radicale et post-dorsale) se dirige vers le palais mou. On peut également parler de voyelles moyennes lorsque la région médiane de la langue (zone dorsale) se dirige vers un point médian de la voûte palatine, mais ce terme est peu utilisé pour le français.

Un autre niveau descriptif est celui du mode d'articulation, qui décrit la façon dont s'effectuent les mouvements de l'appareil phonatoire. Ainsi, le degré d'aperture, qui décrit les mouvements verticaux de la langue, est mesuré pour les voyelles par la distance séparant la langue, à l'endroit où son gonflement est maximal, de la voûte du palais. En français, on distingue ainsi les voyelles fermées et les voyelles mi-fermées, pour lesquelles l'aperture est minimale, à la limite de l'obstruction du canal buccal, puis les voyelles mi-ouvertes et les voyelles ouvertes, pour lesquelles l'aperture est maximale.

En associant l'antériorité (ou lieu d'articulation) et l'aperture, on peut organiser les voyelles dans une disposition trapézoïdale, avec au sommet les voyelles fermées, des antérieures (/i/) jusqu'aux postérieures (/u/), et à la base les voyelles ouvertes (essentiellement /a/). On retrouve aux positions extrêmes les trois voyelles /a/ /i/ /u/, présentes dans la quasi-totalité des langues du monde.

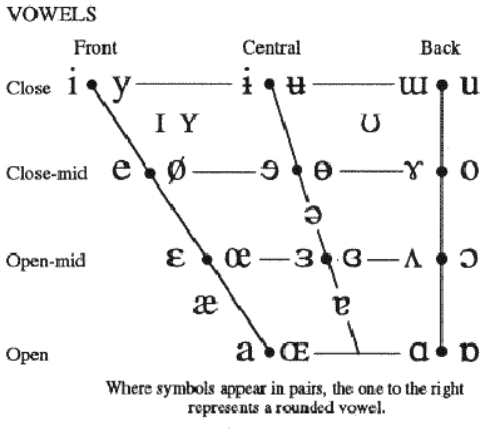


FIG. 2.4 – Voyelles de l'alphabet phonétique international

Toujours en ce qui concerne le mode d'articulation, la labialisation†, ou arrondissement†, désigne le mouvement plus ou moins marqué des lèvres vers l'avant (protrusion) pouvant accompagner la prononciation des voyelles, ce qui permet de distinguer les voyelles labialisées des voyelles non labialisées. Cette distinction permet de mettre en valeur une des originalités du français dans la présence du groupe de voyelles antérieures labialisées /y/ /ø/ /œ/. Ces voyelles sont parfois appelées voyelles orales composées, par opposition aux voyelles orales simples, ce qui rappelle que leur prononciation réunit deux mouvements articulatoires rarement associés : l'articulation antérieure et la labialisation.

La nasalisation† des voyelles se produit lorsque l'on abaisse le voile du palais, ce qui a pour effet de mettre en communication le canal pharyngo-buccal avec les fosses nasales et de

laisser ainsi passer une partie de l'air par le nez. Ainsi, lors de la prononciation des voyelles nasales†, le flux d'air suit à la fois la voie buccale et la voie nasale, tandis qu'il n'emprunte que la voie buccale pour les voyelles orales†. En français, la nasalisation s'accompagne d'une légère compression de l'espace pharyngo-buccal due à l'abaissement prononcé du voile du palais. Cet ajustement des volumes articulatoires permet un équilibrage des résonances des cavités nasale et pharyngo-buccale, ce qui joue un rôle primordial dans le timbre final des voyelles nasales. Le groupe des voyelles nasales du français est /ɑ̃/ /ɔ̃/ /ɛ̃/ /œ̃/.

La durée de prononciation des voyelles permet de faire la distinction entre les voyelles longues et les voyelles brèves. Elle constitue une caractéristique phonétique à part entière puisqu'elle peut introduire des différences de sens entre les mots, en particulier en anglais. En français, cependant, la durée des voyelles a tendance à perdre sa valeur distinctive et de moins en moins de locuteurs différencient des mots comme, par exemple, *mettre* /mɛtr/ et *maître* /mɛ :tr/.

La tension des voyelles, qui oppose les voyelles tendues aux voyelles relâchées, ne peut pas jouer un rôle de structuration du système phonétique en français, puisque toutes les voyelles sont nettement tendues. La prononciation d'une voyelle tendue est caractérisée par une stabilité de la position des articulateurs pendant toute la durée de la vocalisation. Au contraire, au cours de la prononciation d'une voyelle relâchée, les positions des articulateurs suivent une plus grande instabilité, ce qui provoque des changements de timbre pendant la vocalisation. Ainsi, toutes les voyelles de l'anglais sont prononcées de façon naturellement relâchée.

La tension des voyelles françaises constitue une difficulté lors de l'apprentissage du français en tant que langue étrangère ; on observe souvent un trop grand relâchement dans la prononciation, les voyelles fermées /i/ /y/ /u/ ayant tendance à être prononcées comme des semi-fermées /e/ /ø/ /o/, les semi-fermées /e/ /ø/ /o/ comme des semi-ouvertes /ɛ/ /œ/ /ɔ/, et ainsi de suite.

Les diphtongues et les triphtongues sont des voyelles dont la prononciation met successivement en jeu deux ou trois configurations différentes de l'appareil phonatoire. Le son généré est donc une succession de deux ou trois voyelles séparées par des transitions rapides. Les diphtongues et les triphtongues sont particulièrement présentes en anglais, où l'on distingue douze diphtongues fondamentales. Elles sont absentes du français contemporain, bien que subsistant parfois sous des formes orthographiques comme dans *beau*. Ce sont typiquement des voyelles relâchées.

2.2.4 Caractéristiques articulatoires des consonnes

Les consonnes† sont les sons produits par la gêne ou l'arrêt du passage de l'air dans l'espace supralaryngé. A l'inverse des voyelles, la nomenclature phonétique met plus l'accent sur le mode d'articulation en ce qui concerne les consonnes.

Pour les consonnes, le degré d'aperture permet de faire une première distinction. Les consonnes occlusives† /p/ /t/ /k/ /b/ /d/ /g/ (ou consonnes plosives†), sont produites par une fermeture complète du canal buccal. Les consonnes constrictives† /s/ /ʃ/ /f/ /z/ /ʒ/ /v/ (ou consonnes fricatives†), correspondent à un fort resserrement de ce même canal, laissant encore suffisamment d'espace pour permettre l'émission d'un son continu. Enfin, les consonnes sonantes† /m/ /n/ /ɲ/ /l/ /ʀ/ sont associées à un obstacle articulatoire de très faible importance, ce qui les fait se rapprocher nettement du type vocalique.

On peut également mentionner la catégorie des consonnes semi-occlusives†, ou consonnes affriquées†, plus marginale. Ces consonnes sont produites par une succession rapide de deux

consonnes sans changement du lieu d'articulation. On distingue entre autres /ts/, comme dans *tsar*, /tʃ/, comme dans *tchèque*, et /dʒ/, comme dans *jazz*.

Parmi les consonnes sonantes, on distingue les consonnes latérales† comme /l/, pour lesquelles la ligne médiane de la langue est mise en contact avec la voûte palatine, l'air passant par l'espace libre laissé des deux côtés de la langue, et les consonnes vibrantes† comme /ʀ/, résultant de battements rapides accompagnés de brèves résonances, provoqués par une série d'occlusions du canal buccal par les organes articulatoires. En français, on distingue le /r/ apical « roulé », dû à des battements de la pointe de la langue contre les alvéoles et encore utilisé dans de nombreuses régions, du /ʀ/ dorsal, provoqué par le rapprochement vers la luette de la partie postérieure du dos de la langue. Il est souvent articulé comme une fricative, la luette ne vibrant pas dans ce cas.

On utilise également la sonorité pour décrire les consonnes. Elle permet de distinguer les consonnes sourdes, qui sont articulées sans que les cordes vocales entrent en vibration, des consonnes sonores, dont l'articulation s'accompagne au contraire d'une mise en vibration des cordes vocales. Cette différenciation permet d'associer les consonnes par paires, selon qu'elles sont sourdes ou sonores, tous les autres paramètres d'articulation restant identiques par ailleurs : /p/&/b/, /t/&/d/, /k/&/g/, /s/&/z/, /ʃ/&/ʒ/, /f/&/v/.

La nasalisation† indique, comme pour les voyelles, un abaissement du voile du palais et une mise en résonance des fosses nasales. On parle alors de consonnes nasales† ou de consonnes orales†, selon la même distinction faite pour les voyelles. Ainsi les consonnes /m/ /n/ /ɲ/ sont toutes des consonnes occlusives nasales sonores.

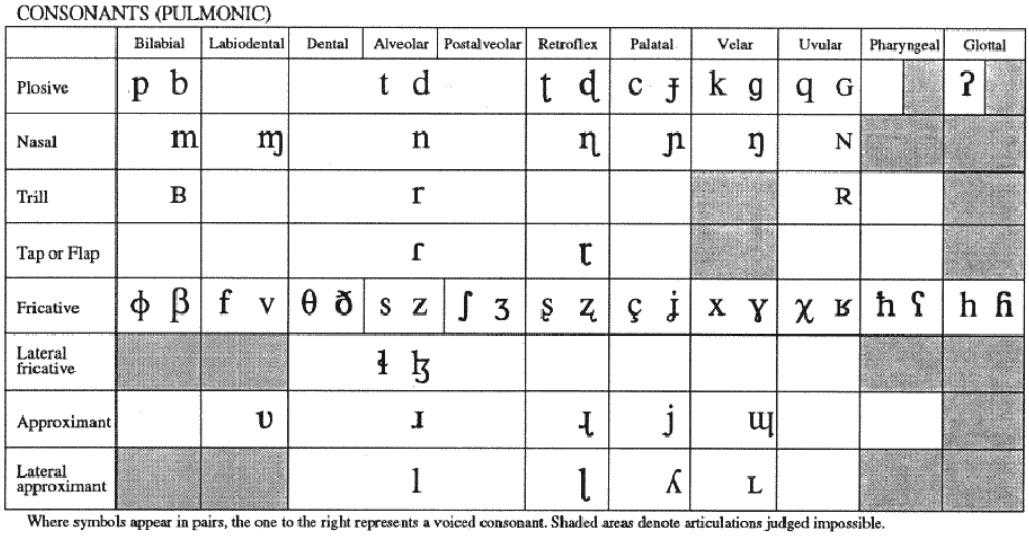
Le *lieu d'articulation* devient un indicateur bien plus précis dans le cas des consonnes. On mentionne le lieu d'articulation à proprement parler, où se produit l'obstruction totale ou partielle du canal buccal, et l'organe d'articulation responsable de cette obstruction.

Mode d'articulation	**Lieu d'articulation**					
	bilabiale	labiodentale	alvéolaire	palatale	vélaire	uvulaire
occlusive sourde	p		t		k	
occlusive sonore	b		d		g	
nasale sonore	m		n	ɲ		
constrictive sourde		f	s	ʃ		ʀ
constrictive sonore		v	z	ʒ		
latérale sonore			l			
glide sonore				j ɥ	w	

Tab. 2.4 – Tableau phonétique des consonnes de la langue française

2.2.5 Les glides

En observant la classification des voyelles suivant leur degré d'aperture, on peut remarquer que pour les voyelles fermées /i/ /y/ /u/, le passage emprunté par l'air en provenance du larynx est relativement étroit. Si l'on continue à élever le dos de la langue vers la voûte palatine, l'articulation entrera dans une zone transitoire où elle donnera naissance à des sons qui ne seront plus totalement vocaliques, sans toutefois aller jusqu'à produire un bruit de constriction. L'ambiguïté résultant de la question de leur classification en tant que voyelles ou en tant que consonnes a conduit à introduire le terme glides† pour désigner ces sons. Le français comporte les


CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	N		
Trill	B			r					R		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

FIG. 2.5 – Consonnes de l'alphabet phonétique international

trois glides /j/ /ɥ/ /w/, respectivement présents dans les mots *ciel* /sjɛl/, *nuit* /nɥi/ et *Louis* /lwi/.

2.3 Caractéristiques acoustiques des phonèmes[3]

2.3.1 Structure acoustique des phonèmes

L'étude des mécanismes articulatoires utilisés par l'être humain pour produire de la parole montre que l'on peut répartir les sons du langage en deux types. Les sons voisés, comme les voyelles, sont produits par une mise en vibration des cordes vocales qui génère des fluctuations périodiques de pression acoustique. Les sons non voisés, comme les consonnes sourdes, sont produits par des constrictions ou des occlusions du flux d'air issu des poumons, ce qui génère des fluctuations apériodiques, équivalent à du bruit acoustique. Pour certains sons, comme les consonnes sonores, on obtient un couplage de fluctuations périodiques et de fluctuations apériodiques, le locuteur associant un voisement et une occlusion ou une constriction.

Lorsqu'un son voisé est émis, les cordes vocales entrent en vibration à une fréquence particulière, essentiellement déterminée par les caractéristiques mécaniques des cordes vocales et par la pression subglottique. Cette fréquence, la fréquence fondamentale†, présente une valeur moyenne de 120 Hz chez l'homme et de 220 Hz chez la femme. Elle correspond à la fréquence de la composante spectrale la plus basse de l'onde acoustique émise par le larynx.

Après son émission, le son laryngé traverse les différentes cavités de l'espace supralaryngé. On y trouve les cavités pharyngale, buccale postérieure et buccale antérieure, labiale et nasale, ces différents volumes pouvant être déformés et mis en communication par les organes de la phonation. Chacune de ces cavités va jouer le rôle d'un résonateur acoustique et va amplifier les fréquences spectrales des sons laryngés proches de sa fréquence propre de résonance. L'articulation permet de contrôler et de modifier l'ensemble de ces fréquences de résonance, et donc de déterminer les plages fréquentielles qui seront amplifiées sélectivement.

[3] d'après Autesserre (1991); Chiss *et al.* (1996)

Par analyse acoustique de la parole, on peut mettre en évidence ce phénomène d'amplification sélective. Cette analyse est réalisée au moyen d'un spectrogramme†, qui est une représentation des variations de l'intensité des ondes sonores de la parole en fonction du temps écoulé et de la fréquence considérée. Usuellement, le temps se lit horizontalement et la fréquence verticalement, la coloration du spectrogramme étant mise en correspondance avec une échelle d'intensité acoustique.

Les différentes résonances des cavités supralaryngées se traduisent par des accumulations particulières d'énergie acoustique au voisinage de certaines fréquences, correspondant à des bandes sombres sur le spectrogramme. Selon que leur fréquence centrale croît, décroît ou reste stable au cours du temps, ces bandes peuvent prendre une orientation horizontale, montante ou descendante. Ces pics fréquentiels particuliers sont appelés des formants†, et sont des caractéristiques essentielles permettant de décrire la structure acoustique des phonèmes.

On numérote usuellement les formants en fonction de leur fréquence centrale : F1, F2, F3. Les fréquences couvertes par la première zone formantielle (F1) peuvent varier approximativement de 200 à 900 Hz. Le deuxième formant F2, plus haut, correspond à des fréquences variant approximativement entre 500 et 2500 Hz. Les rapports entre les cavités de résonance mises en jeu et la répartition fréquentielle des formants sont complexes. En simplifiant grossièrement, on peut associer le premier formant F1 à la cavité pharyngale et à la cavité buccale postérieure, tandis que le second formant F2 peut être associé à la cavité buccale antérieure et à la cavité labiale.

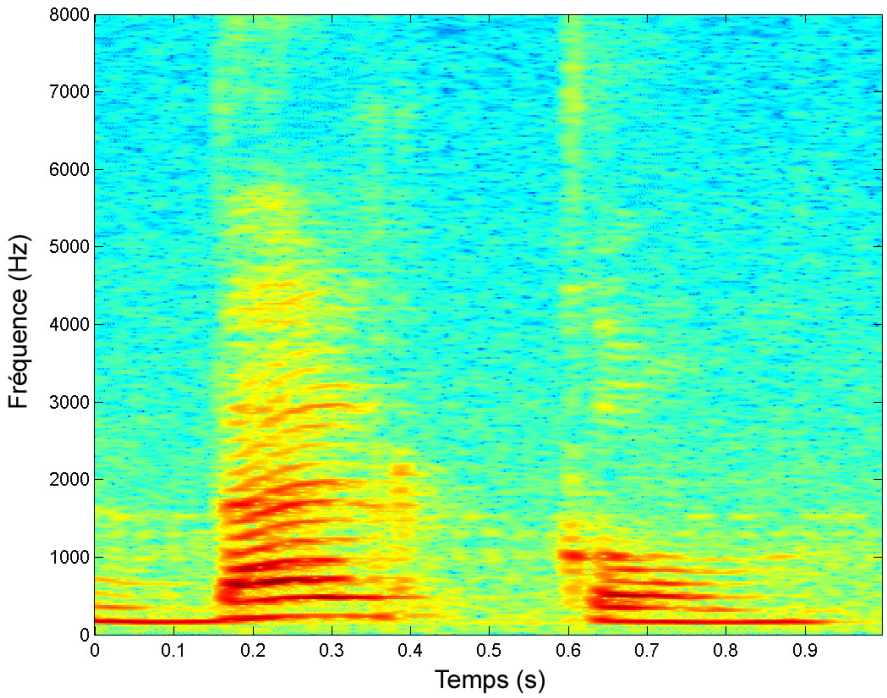


FIG. 2.6 – Exemple de spectrogramme pour la prononciation du mot « balcon ». On peut nettement distinguer les deux syllabes du mot, ainsi que le trait de voisement au début de la prononciation, qui précède le phonème /b/.

2.3.2 Caractéristiques acoustiques des voyelles

Pendant la production articulatoire d'une voyelle, la forme du conduit vocal se modifie très peu. Ainsi la fréquence des formants associée à une voyelle particulière est stable d'une production à l'autre, aussi bien au niveau intra- qu'au niveau inter-individuel. Pour chaque voyelle on peut donc isoler plusieurs formants qui lui sont spécifiques. On constate en pratique que les formants F1 et F2, les deux plus bas dans l'échelle des fréquences, sont les plus significatifs pour l'identification acoustique des sons. Le formant F3 intervient parfois également comme composante acoustique discriminatoire. Les formants les plus hauts, à partir de F4, jouent un rôle plus secondaire et servent essentiellement à caractériser le timbre de la voix.

Sur l'échelle des fréquences, les voyelles antérieures sont caractérisées par des formants F1 et F2 nettement distincts (et par un formant F3 proche de F2). Les voyelles postérieures présentent au contraire des formants F1 et F2 proches l'un de l'autre. Ces caractéristiques acoustiques permettent de distinguer deux types vocaliques : les voyelles diffuses, dont les deux premiers formants occupent des domaines fréquentiels éloignés, et les voyelles compactes, dont les deux premiers formants occupent des domaines fréquentiels proches. Plus les formants F1 et F2 sont éloignés, et plus il est facile de discriminer leurs sources acoustiques en attribuant F1 à la résonance de la cavité buccale postérieure et F2 à la résonance de la cavité buccale postérieure. Au contraire, plus ils sont proches, moins cette distinction entre les différentes sources acoustiques est évidente.

On place usuellement les voyelles sur un diagramme acoustique où F1 est représenté sur l'axe des ordonnées, par fréquences décroissantes, et F2 sur l'axe des abscisses, également par fréquences décroissantes. Ce type de diagramme permet de montrer l'association entre la phonétique acoustique et la phonétique articulatoire. Lorsque l'on passe progressivement de /i/ à /a/, on observe une augmentation de la fréquence de F1 et une diminution de la fréquence de F2 par élévation progressive du degré d'aperture. Lorsque l'on passe de /i/ à /y/, on peut mettre en parallèle la labialisation avec l'abaissement de la fréquence de F2, par agrandissement du volume de la cavité labiale.

Les voyelles nasales peuvent être caractérisées par un formant F1 d'intensité légèrement inférieure à celle des voyelles orales correspondantes, ainsi que par un formant FN1 centré sur une fréquence relativement stable de 250 Hz. Ce formant est dû à la résonance acoustique des cavités nasales, toutefois il ne joue qu'un rôle secondaire pour l'identification acoustique des voyelles nasales.

2.3.3 Caractéristiques acoustiques des consonnes

Les consonnes peuvent être caractérisées acoustiquement par les formants de bruit et les transitions formantiques. Les formants de bruit sont des bandes de fréquence relativement étendues sur lesquelles les intensités spectrales sont particulièrement élevées. Ces formants correspondent aux bruits d'explosion et de constriction générés lors de l'articulation des consonnes. Les transitions formantiques sont des modifications apportées par les consonnes aux voyelles qui les précédent et les suivent. Ces modifications se présentent sous la forme de glissements des fréquences formantielles des voyelles articulées avant et après la consonne concernée. Elles interviennent dans l'étude des mécanismes de la coarticulation.

Les transitions formantiques jouent un rôle essentiel dans la caractérisation acoustique des consonnes occlusives. Leur structure acoustique peut se décomposer en deux étapes temporelles successives. On observe dans un premier temps une transition formantique, avec des changements rapides dans les fréquences des formants vocaliques qui précèdent la consonne occlusive. Cette

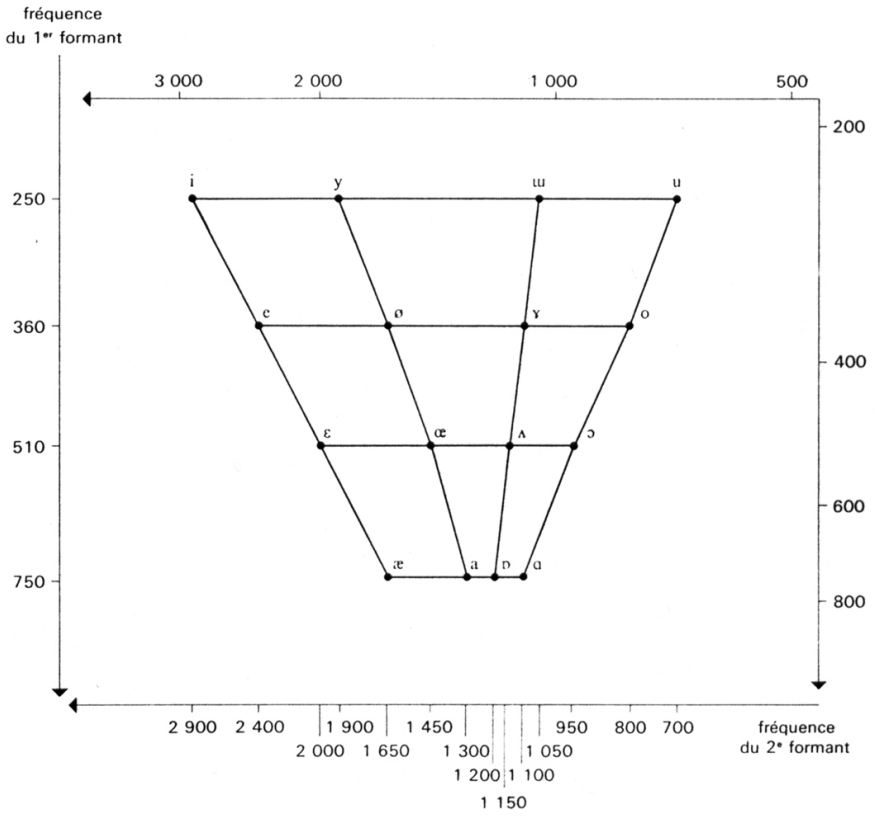


FIG. 2.7 – Diagramme acoustique des voyelles orales du français. D'après Autesserre (1991).

première étape est associée à l'obstruction du canal buccal par les articulateurs. Puis on observe un formant de bruit correspondant au bruit d'explosion généré par la séparation brusque des articulateurs, qui marque la fin de l'obstruction du canal buccal.

On a constaté que la reproduction artificielle des transitions formantiques seules, sans bruit d'occlusion, suffisait pour permettre une identification correcte de la consonne occlusive associée à ces transitions. On peut donc discriminer les consonnes occlusives à partir de ces transitions et des fréquences vers lesquelles elles convergent. On appelle notamment locus d'une consonne la fréquence vers laquelle convergent les transitions de deuxième formant associées à cette consonne.

Les occlusives bilabiales /p/ /b/ ont un locus dans le domaine des basses fréquences, aux environs de 700 Hz ; les occlusives labiodentales /t/ /d/ ont un locus nettement plus élevé, aux environs de 1800 Hz ; les occlusives vélaires /k/ /g/ ont quant à elles un locus très variable selon les voyelles qui les accompagnent. Les occlusives nasales /m/ /n/ /ɲ/ sont caractérisées par un locus identique aux occlusives orales /b/ /d/ qui leur sont associées, auquel se rejoute un formant nasal à 250 Hz.

Les formants de bruit jouent un rôle plus important dans la caractérisation acoustique des consonnes constrictives. Les paires /s/ /z/ et /ʃ/ /ʒ/ se distinguent les unes des autres par la bande de fréquence couverte par leurs formants de bruit. Cette bande s'étend jusqu'à 4000 Hz pour les sifflantes /s/ /z/, tandis qu'elle ne monte qu'à 2000 Hz pour les chuintantes /ʃ/ /ʒ/. Le formant de bruit de la paire /f/ /v/ est inégalement réparti dans les hautes fréquences, on identifie plutôt les consonnes de cette paire par leur locus, similaire à celui des occlusives

bilabiales.

Les sonantes /l/ /ʀ/ ont une structure spectrale très proche de celle des voyelles, on leur donne pour cela le nom de consonnes liquides[†]. Ces consonnes sont caractérisées par des formants d'intensité plus faible que pour les voyelles, par un formant F1 d'une fréquence avoisinant les 400 Hz et par un formant de bruit résiduel. Leurs transitions formantiques assurent une continuité entre leurs formants et ceux des voyelles qui les accompagnent. Les glides /j/ /ɥ/ /w/ présentent des caractéristiques similaires mais leurs formants se distinguent par une durée plus brève.

Bibliographie

Abitbol, J. (1991). Phonation. *In Encyclopaedia universalis*, volume 18, pages 91–95. Encyclopaedia universalis.

Autesserre, M. (1991). Phonétique. *In Encyclopaedia universalis*, volume 18, pages 95–100. Encyclopaedia universalis.

Chiss, J.-L., Filliolet, J. et Maingueneau, D. (1996). *Linguistique française*. Hachette.

Chapitre 3

Perception audiovisuelle de la parole

Sommaire

3.1 Développement de l'intégration audiovisuelle de la parole 60
3.1.1 Reconnaissance de la voix chez les enfants 60
3.1.2 Reconnaissance du visage chez les enfants 60
3.1.3 Perception audiovisuelle de la parole chez les enfants 60
3.2 Facilitations perceptives par la parole audiovisuelle 61
3.2.1 Améliorations perceptives en condition non dégradée 61
3.2.2 Améliorations perceptives en condition dégradée 61
3.3 Facteurs influant sur l'intégration audiovisuelle de la parole 62
3.3.1 Congruence temporelle de la parole 62
3.3.2 Congruence phonétique de la parole : complémentarité audiovisuelle . . 63
3.3.3 Facteurs linguistiques 64
3.4 Conséquences d'une incongruence spatio-temporelle 64
3.4.1 Incongruence temporelle de la parole 64
3.4.2 Incongruence spatiale de la parole 65
3.5 Conséquences d'une incongruence phonétique : effet McGurk 65
3.5.1 Principe 65
3.5.2 Facteurs perceptifs 66
3.5.3 Facteurs linguistiques 67
3.6 Mécanismes d'intégration audiovisuelle de la parole 67
3.6.1 Modèles d'intégration audiovisuelle tardive 67
3.6.2 Modèles d'intégration audiovisuelle précoce 68
3.6.3 Intégration perceptive audio-visuo-motrice de la parole 69
3.6.4 Différence d'intégration audiovisuelle entre hommes et femmes 71
Bibliographie 73

Introduction

A chaque fois que la parole est accessible simultanément par les voies auditive et visuelle, le cerveau réalise une intégration multimodale des signaux auditifs et visuels pour aboutir à un décodage des signaux de parole sous-jacents. Cette coopération audiovisuelle

est une caractéristique essentielle de la perception de la parole. Elle intervient dès les premières semaines de la vie extra-utérine, en contribuant à la maturation des réseaux multisensoriels impliqués dans la parole et le langage. Pour chaque situation de perception audiovisuelle de la parole, ces réseaux sont systématiquement mis à contribution en permettant une meilleure isolation et une meilleure discrimination des informations pertinentes, ceci grâce à un traitement complexe des indices auditifs, des indices visuels et des congruences audiovisuelles reliant ces deux modalités perceptives.

Dans le cas où les signaux optiques et acoustiques sont présents simultanément, comme pour la communication en face-à-face, les indices associés aux différentes modalités perceptives sont couplés au niveau neuronal pour aboutir à un percept unifié : c'est l'intégration audiovisuelle, mécanisme fondamental pour la perception audiovisuelle de la parole. Dans ce contexte nous nous attarderons en particulier sur l'effet McGurk, qui fournit un moyen intéressant et original de mettre en lumière les processus perceptifs qui sous-tendent l'intégration polymodale de la parole. Cette synthèse entre les modalités auditive et visuelle n'est par ailleurs qu'un cas particulier de la notion plus générale d'intégration multisensorielle, qui sera abordée de façon détaillée dans la suite de notre exposé (« *Intégration multisensorielle* », ch. 5, p. 109).

3.1 Développement de l'intégration audiovisuelle de la parole

3.1.1 Reconnaissance de la voix chez les enfants

Dès les premiers stades du développement de l'être humain, on observe l'émergence de compétences spécialisées dans l'identification et la reconnaissance de la voix. Ainsi, vers la fin de la gestation, le fœtus humain est attentif aux sons et aux voix qui lui parviennent à travers la paroi de l'utérus (Birnholtz et Benacerraf, 1983; Querleu *et al.*, 1989). Au stade prénatal, on observe un apprentissage auditif de la voix maternelle et de la prosodie de la parole (DeCasper et Fifer, 1980; DeCasper et Spence, 1986). Les nouveaux-nés sont capables de discriminer des voix du même sexe (DeCasper et Prescott, 1984; Bahrick *et al.*, 2003) et de discriminer les sexes en fonction de la voix (Miller *et al.*, 1982; Miller, 1983).

3.1.2 Reconnaissance du visage chez les enfants

Très tôt au cours du développement humain, on observe l'apparition de compétences perceptives spécfiques à la reconnaissance des visages humains. Les nouveaux-nés montrent une nette préférence pour les visages (Easterbrook *et al.*, 1999; Maurer et Barrera, 1981) et savent discriminer facilement les visages dès les premiers mois de la vie (Fagan, 1972; Kleiner, 1987; Maurer et Young, 1983; Pascalis et de Schonen, 1994). Ils sont rapidement capables de faire des distinctions entre les visages, qu'ils soient familiers ou non familiers (Barrera et Maurer, 1981; Bushnell, 1982; Walton *et al.*, 1992; Pascalis *et al.*, 1995, 1998).

3.1.3 Perception audiovisuelle de la parole chez les enfants

Dès les premières semaines, on peut observer chez les jeunes enfants une préférence marquée pour les signaux de parole audiovisuelle dont la congruence temporelle et spatiale est respectée. Ainsi les jeunes enfants sont particulièrement sensibles à la congruence spatiale de la parole : dans des paradigmes expérimentaux où l'on introduit un décalage spatial entre le visage et la

composante sonore de la voix, des enfants de 1 à 2 mois se montrent stressés (Aronson et Rosenbloom, 1971; Spelke et Owsley, 1979). On observe également une sensibilité pour la congruence temporelle : lorsqu'on introduit un déphasage temporel entre la composante sonore de la voix et les mouvements du visage et des lèvres, des enfants de 3 à 4 mois se montrent nettement moins attentifs que pour un signal audiovisuel intact (Dodd, 1979; Pickens *et al.*, 1994; Spelke et Cortelyou, 1980; Walker, 1982).

La sensibilité des enfants pour la congruence audiovisuelle de la parole s'étend au-delà du simple niveau de la congruence spatio-temporelle : en effet on remarque également une sensibilité pour la correspondance phonétique entre les signaux visuels et auditifs (Kuhl et Meltzoff, 1982, 1984, 1988; MacKain *et al.*, 1983; Walton et Bower, 1993; Legerstee, 1990; Patterson et Werker, 1999, 2002, 2003). Entre 5 et 7 mois, les jeunes enfants sont capables d'établir des correspondances entre les visages et les voix à partir des émotions qui sont exprimées (Soken et Pick, 1992; Walker, 1982; Walker-Andrews, 1997), à partir du sexe (Walker-Andrews et Lennon, 1991) et également à partir de l'âge (Bahrick *et al.*, 1998).

L'importance d'une stimulation audiovisuelle pour l'acquisition de la parole est démontrée par les observations réalisées chez des enfants déficients visuels ou auditifs, pour lesquels la coopération multisensorielle se trouve entravée. Ainsi, on observe que les enfants aveugles rencontrent des difficultés pour apprendre certains contrastes phonétiques particuliers comme le contraste /m/-/n/, qui est peu marqué au travers des indices auditifs mais fortement marqué au travers des indices visuels (Mulford, 1988).

3.2 Facilitations perceptives par la parole audiovisuelle

3.2.1 Améliorations perceptives en condition non dégradée

Lorsque la composante auditive de la parole n'est pas dégradée par des bruits parasites ou par un dysfonctionnement du système auditif, on montre que les indices visuels spécifiques de la parole (mouvements des lèvres et du visage) permettent d'abaisser le seuil de détection auditive de la parole de plusieurs décibels (Grant et Seitz, 2000). Dans des conditions de préservation de la composante auditive de la parole, ces indices visuels permettent également une meilleure perception de la parole lorsqu'elle est prononcée dans une langue étrangère non familière, lorsqu'elle est prononcée dans la langue natale mais avec un accent étranger, et également lorsque la parole est parfaitement intelligible mais caractérisée par haut niveau de complexité sémantique ou syntaxique (Reisberg *et al.*, 1987; Arnold et Hill, 2001; Kim et Davis, 2003).

3.2.2 Améliorations perceptives en condition dégradée

De nombreuses expériences ont utilisé de la parole acoustiquement dégradée, en comparant les performances perceptives en conditions auditive seule et audiovisuelle. Toutes ces expériences concordent pour démontrer l'avantage perceptif apporté par les informations visuelles en complément des informations auditives dégradées (Cotton, 1935; Sumby et Pollack, 1954; Miller et Nicely, 1955; Neely, 1956; Erber, 1969, 1975; Binnie *et al.*, 1974; Summerfield, 1979; MacLeod et Summerfield, 1987; Grant et Braida, 1991; Benoit *et al.*, 1994). Nous soulignons en particulier les expériences pionnières réalisées par Cotton (1935), qui constitue probablement la première mise en évidence scientifique de l'avantage perceptif apporté par la parole audiovisuelle, et par Sumby et Pollack (1954), qui sont parmi les premiers à avoir quantifié cet avantage perceptif audiovisuel.

L'apport des informations visuelles n'est toutefois pas illimité, puisqu'il disparaît lorsque le

niveau de dégradation acoustique dépasse un certain seuil. Il existe donc un niveau de dégradation intermédiaire pour lequel le gain audiovisuel est maximal (voir à ce sujet *« Multisensorialité : principes et propriétés »*, sect. 5.1, p. 110). Dans le cas où la parole est dégradée par du bruit blanc, on observe expérimentalement un gain audiovisuel maximal pour des rapports signal/bruit voisins de -12 dB (Ross *et al.*, 2007).

Ces expériences réalisées par dégradation artificielle du signal acoustique de parole chez le sujet normo-entendant sont complétées par un corpus volumineux d'études réalisées chez les patients sourds, qui montrent elles aussi l'amélioration perceptive fournie par la parole audiovisuelle dans le cas où le signal auditif n'est pas correctement restitué (voir *« Perception audiovisuelle de la parole lors de la surdité »*, § 6.4.3, p. 138).

3.3 Facteurs influant sur l'intégration audiovisuelle de la parole

3.3.1 Congruence temporelle de la parole

La parole audiovisuelle ne résulte pas d'une juxtaposition arbitraire d'informations auditives et visuelles ; au contraire il existe une congruence temporelle et spatiale très forte entre ces deux composantes. Dans une étude visant à analyser de façon fine l'effet de la corrélation des signaux auditifs et visuels (Grant et Seitz, 2000), les auteurs démontrent que l'amélioration de la détectabilité audiovisuelle est liée à la corrélation entre les fluctuations de l'amplitude acoustique et les mouvements d'ouverture de la bouche. Une étude ultérieure (Grant, 2001) a montré que l'amélioration audiovisuelle est maximale dans la bande de fréquence correspondant aux deuxième et troisième formants. Or ce domaine fréquentiel correspond précisément à celui où la corrélation temporelle audiovisuelle de la parole est maximale, ce qui suggère que la corrélation temporelle audiovisuelle de la parole est directement impliquée dans l'amélioration perceptive audiovisuelle. Cette amélioration persiste même lorsque les participants sont stimulés dans une langue étrangère inconnue (Kim et Davis, 2003), ce qui renforce l'hypothèse d'une influence directe de la corrélation temporelle audiovisuelle sur le gain perceptif audiovisuel.

Afin de distinguer les rôles respectifs de la congruence temporelle et de la congruence phonétique (voir ci-après) dans l'intégration de la parole audiovisuelle, il est possible de concevoir des stimuli dont la composante visuelle est toujours la même : cette dernière n'apporte alors aucune information phonétique permettant de discriminer les stimuli. L'utilisation de ce type de protocole a permis de démontrer que la congruence temporelle audiovisuelle prise isolément, en excluant volontairement les indices fournis par la congruence phonétique audiovisuelle, suffit à produire une amélioration de l'intelligibilité de la parole dans le bruit (Schwartz *et al.*, 2004).

Une hypothèse possible pour expliquer cette influence serait que l'interlocuteur tire parti de la légère avance — quelques dizaines de millisecondes — du moment d'ouverture maximale de la bouche sur le pic d'énergie dans la bande spectrale correspondant aux deuxième et troisième formants. Une façon de tester cette hypothèse est d'utiliser des signaux de parole inversés temporellement, ce qui présente l'avantage de supprimer l'avance visuelle tout en préservant la corrélation temporelle audiovisuelle (Kim et Davis, 2004). Dans ce cas on n'observe pas d'amélioration perceptive audiovisuelle ; toutefois la réintroduction artificielle d'une avance du signal visuel n'induit pas non plus d'amélioration audiovisuelle, ce qui tend à infirmer l'explication de la facilitation audiovisuelle à partir de l'avance du signal visuel sur le signal auditif.

3.3.2 Congruence phonétique de la parole : complémentarité audiovisuelle

Un des facteurs importants permettant d'expliquer l'amélioration perceptive produite par la présentation audiovisuelle de la parole réside dans la complémentarité naturelle existant entre les informations auditives et visuelles du signal de parole. On pourrait résumer cette complémentarité de façon simplifiée en écrivant que ce sont les phonèmes les moins faciles à discriminer visuellement qui sont les mieux discriminables auditivement, et inversement. Nous aller préciser cette affirmation dans les cas distincts des consonnes et des voyelles.

Nous avons vu dans *« Caractéristiques articulatoires des consonnes »*, § 2.2.4, p. 50 que les consonnes pouvaient être caractérisées par leur mode d'articulation (occlusives, constrictives ou liquides), leur voisement (sourdes ou sonores), leur nasalité et leur lieu d'articulation. Le voisement et la nasalité sont des caractéristiques articulatoires facilement identifiables par les informations auditives, même en condition de dégradation auditive, mais impossibles à discriminer par les informations visuelles. Le lieu d'articulation, au contraire, est facile à identifier visuellement mais très sensible à une dégradation auditive. Enfin le mode d'articulation possède un statut intermédiaire, assez facilement identifiable par l'audition et partiellement visible (Binnie *et al.*, 1974).

La reconnaissance des consonnes en conditions visuelle seule et auditive seule a été testée en particulier dans l'étude de Summerfield (1987). L'arbre de confusion des réponses auditives montre une organisation globalement inverse de son équivalent visuel : ce qui est bien perçu auditivement ne l'est pas visuellement et vice versa. Les résultats expérimentaux montrent notamment une meilleure discrimination visuelle entre les phonèmes /p/, /t/ et /k/ qu'en condition auditive. Inversement, une forte confusion visuelle entre les phonèmes /p/, /b/ et /m/, tous trois caractérisés par le même lieu d'articulation bilabial, disparaît en condition auditive. Walden *et al.* (1977) ont rapporté des résultats similaires dans le cas de sujets spécialement entraînés à la lecture labiale.

Les voyelles peuvent quant à elles être caractérisées par une complémentarité audiovisuelle similaire à celle des consonnes, les voyelles les plus difficiles à discriminer en condition de dégradation auditive étant les plus faciles à identifier visuellement (Benoit *et al.*, 1994). Cette complémentarité se retrouve entre l'arrondissement, l'aperture et l'antériorité, qui sont les principaux traits articulatoires structurant l'espace acoustique des voyelles (Robert-Ribes *et al.*, 1998) (voir *« Caractéristiques articulatoires des voyelles »*, § 2.2.3, p. 48).

Une proposition d'explication de la complémentarité audition/vision (Summerfield et Assmann, 1989) associe les articulateurs visibles (lèvres, dents et partie antérieure de la langue) à la production de sons de fréquence élevée, associés à des fluctuations temporelles rapides (comme lors de l'articulation des consonnes occlusives). Ainsi ces sons seraient caractérisés acoustiquement par des turbulences de faible intensité sonore, facilement dégradées en présence de bruit, ce désavantage auditif pouvant être compensé par la bonne disponibilité des informations visuelles. A l'inverse, les articulateurs non visibles (partie postérieure de la langue, voile du palais, larynx) produiraient des sons constants, de forte intensité, à des fréquences basses caractéristiques notamment du voisement et de la nasalité, et sont donc facilement identifiables par l'audition. Cette explication vient compléter les résultats présentés par Fant (1973), qui associe la résonance des cavités postérieures de l'espace supralaryngé (non visibles) au premier formant acoustique, alors que le second formant correspondrait plutôt à la résonance des cavités antérieures. Ainsi, tandis que le premier formant présente une bonne stabilité acoustique qui rend son identification facile par l'audition, le second présente davantage de variabilité et bénéficie du renfort perceptif

apporté par la vision des mouvements des lèvres et de la langue. Ces hypothèses ont été testées de façon fine dans des études ultérieures (Grant et Walden, 1996; Grant, 2001), qui ont démontré que le gain audiovisuel est le plus important dans la bande de basses fréquences correspondant au premier formant, ces basses fréquences étant caractéristiques des traits de voisement et de nasalité, impossibles à identifier par la vision. Inversement, le gain audiovisuel est le moins important dans la bande de hautes fréquences correspondant aux deuxième et troisième formant, ces hautes fréquences étant caractéristiques du lieu d'articulation (en particulier les lieux antérieurs), facile à identifier par la vision.

3.3.3 Facteurs linguistiques

En plus de la congruence spatio-temporelle et phonétique, l'intégration des composantes auditive et visuelle de la parole peut également être facilitée par des indices de nature linguistique. Ces indices incluent en particulier les indices lexicaux, la perception des mots étant contrainte par la nécessité d'une correspondance avec des mots faisant partie du vocabulaire de l'interlocuteur. Ainsi l'amélioration perceptive audiovisuelle est moins importante, même si elle existe, lorsque les mots sont prononcés dans une langue inconnue de l'interlocuteur (Kim et Davis, 2003). La contrainte de la cohérence lexicale des mots perçus pourrait donc jouer un rôle important lors de l'intégration audiovisuelle de la parole, possiblement par l'intermédiaire de mécanismes cognitifs de type *top-down*. Cette hypothèse est soutenue par l'existence d'effets de listes, la réduction du nombres de réponses possibles induisant une amélioration perceptive audiovisuelle (Sumby et Pollack, 1954). Pour finir, les indices linguistiques peuvent également provenir d'indices cognitifs de plus haut niveau comme des indices syntaxiques ou sémantiques, fortement dépendants du contexte de parole dans lequel le mot est prononcé.

3.4 Conséquences d'une incongruence spatio-temporelle

3.4.1 Incongruence temporelle de la parole

Les effets d'un décalage temporel entre les composantes auditive et visuelle de la parole ont à l'origine été étudiés en vue d'application prothétiques, afin d'évaluer les conséquences de l'introduction d'un délai au cours du traitement du signal acoustique réalisé par la prothèse auditive (McGrath et Summerfield, 1985; Pandey *et al.*, 1986). Les expériences visant à mesurer le seuil temporel de détection de l'asynchronie audiovisuelle indiquent des résultats variables en fonction de la méthodologie de mesure et des stimuli utilisés (Hirsh et Sherrick, 1961; Dixon et Spitz, 1980; McGrath et Summerfield, 1985; Conrey et Pisoni, 2006; Vatakis et Spence, 2006a,b), toutefois elles tendent globalement à indiquer une sensibilité à l'asynchronie plus importante pour les stimuli les plus simples. Par ailleurs toutes ces études concordent pour indiquer une tolérance plus importante pour les retards auditifs que pour les retards visuels. Cette sensibilité temporelle asymétrique peut être expliquée par la différence entre les vitesses de propagation de la lumière et du son, qui provoque un retard auditif naturel (voir à ce sujet « *Congruence temporelle* », § 5.2.2, p. 112). Ainsi la sensation d'incongruence temporelle de la parole se produirait pour des retards visuels plus faibles (Cathiard et Tiberghien, 1994).

Les estimations de la fenêtre d'intégration audiovisuelle temporelle de la parole ont été réalisées au travers de protocoles testant l'amélioration audiovisuelle de l'intelligibilité (McGrath et Summerfield, 1985; Grant *et al.*, 2004) ou l'effet McGurk (Pandey *et al.*, 1986; Munhall *et al.*, 1996; Massaro *et al.*, 1996; Jones et Jarick, 2006; van Wassenhove *et al.*, 2007). Ces estimations s'étendent approximativement jusqu'à 60 ms pour le retard visuel et 240 ms pour le retard auditif,

bien que des effets d'intégration audiovisuelle puissent encore être observés au-delà de ces limites.

3.4.2 Incongruence spatiale de la parole

L'intégration audiovisuelle de la parole a été testée de façon beaucoup plus étendue en cas d'incongruence temporelle qu'en cas d'incongruence spatiale. Les études réalisées dans ce dernier cas ont essentiellement porté sur l'effet McGurk. Elles indiquent une fenêtre d'intégration audiovisuelle spatiale s'étendant jusqu'à un décalage angulaire de 180° pour la composante auditive, la composante visuelle restant face au sujet (Jones et Munhall, 1997; Jones et Jarick, 2006), et jusqu'à un décalage de 60° pour la composante visuelle, la composante auditive restant face au sujet (Paré *et al.*, 2003). Cependant le décalage visuel s'accompagne d'une diminution des percepts de type McGurk, cette diminution pouvant être expliquée par la diminution de la résolution spatiale du système visuel en champ périphérique.

3.5 Conséquences d'une incongruence phonétique : effet McGurk

3.5.1 Principe

Si l'intégration de stimuli polymodaux congruents engendre des améliorations perceptives (voir « *Intégration multisensorielle* », ch. 5, p. 109), elle peut également les altérer et créer des altérations perceptives dans le cas ou la congruence spatiale ou temporelle n'est plus respectée (voir notamment « *Multisensorialité et congruence spatio-temporelle* », sect. 5.2, p. 110). Dans le cas particulier de la parole audiovisuelle, il est possible d'aller plus loin et de préserver la congruence spatiale et temporelle des signaux de parole auditifs et visuels tout en y introduisant une incongruence phonétique : on provoque alors une illusion perceptive appelée effet McGurk (McGurk et MacDonald, 1976; MacDonald et McGurk, 1978).

L'effet McGurk montre que des informations visuelles articulatoires portant sur les mouvements des lèvres sont intégrées inconsciemment dans notre perception auditive et sont susceptibles d'influencer notre compréhension de la parole. Cet effet se produit lorsqu'un stimulus auditif représentant un phonème particulier est doublé par un stimulus visuel représentant un phonème différent. Le percept résultant est souvent un phonème différent de celui présenté dans le stimulus auditif (en particulier lorsque les lieux d'articulation auditif et visuel sont différents), ce qui met en évidence une intervention de la vision dans la catégorisation phonétique.

Par exemple, des auditeurs perçoivent typiquement le phonème /da/ lorsqu'on leur présente simultanément les phonèmes auditif /ba/ et visuel /ga/, bien que le phonème auditif soit clairement identifié comme /ba/ lorsqu'il est présenté sans contrepartie visuelle incongruente (McGurk et MacDonald, 1976; MacDonald et McGurk, 1978). Ce phénomène indique que l'information contenue dans le signal acoustique est intégrée perceptuellement avec l'information contenue dans le signal optique pour parvenir à une interprétation phonétique unifiée.

Dans le cadre d'une incongruence audiovisuelle phonétique comme l'effet McGurk, il est d'usage d'établir une distinction entre quatre classes possibles de percepts :

- la perception du même lieu d'articulation que pour le phonème auditif ; on parle alors de *dominance auditive* (ce cas correspond à une absence d'effet McGurk). Par exemple, un /ba/ auditif avec un /da/ visuel aboutissant à la perception d'un /ba/ ;
- la perception du même lieu d'articulation que pour le phonème visuel ; on parle alors de *dominance visuelle*. Par exemple, un /ba/ auditif avec un /da/ visuel aboutissant à la perception d'un /da/ ;

- la perception d'un lieu d'articulation intermédiaire, n'existant ni dans le phonème auditif ni dans le phonème visuel ; on parle de percept *intermédiaire* ou *fusionné*. Par exemple, un /ba/ auditif avec un /ga/ visuel aboutissant à la perception d'un /da/. Toutefois, si /g/ et /d/ ne sont pas considérés comme discriminables visuellement, on se retrouve encore dans un cas de dominance visuelle ;
- la perception des deux lieux d'articulation ; on a alors une *combinaison*. Par exemple, un /da/ auditif et un /ba/ visuel aboutissant à la perception d'un /bda/.

3.5.2 Facteurs perceptifs

L'effet McGurk a pu être répliqué dans différentes langues (Massaro et Ferguson, 1993; Sekiyama et Tokhura, 1993), lorsque les composantes visuelle et auditive proviennent de locuteurs de sexes différents (Green et Kuhl, 1991) et même dans une condition audio-tactile, lorsque les sujets touchent le visage du locuteur tout en écoutant sa voix (Fowler et Dekle, 1991). Il existe également un effet McGurk en cas d'incongruence phonétique audiovisuelle pour les voyelles (Summerfield et McGrath, 1984; Lisker et Rossi, 1992), bien qu'il soit moins marqué que pour les consonnes. Il est intéressant de remarquer que cet effet fonctionne également chez les jeunes enfants (McGurk et MacDonald, 1976; Desjardins *et al.*, 1997; Rosenblum *et al.*, 1997), bien qu'il soit moins marqué chez ces derniers que chez les adultes (McGurk et MacDonald, 1976; Desjardins *et al.*, 1997).

D'un point de vue phénoménologique, on peut faire plusieurs observations intéressantes sur l'effet McGurk. Très souvent, les sujets ne se rendent pas compte de la non-congruence audiovisuelle du stimulus (Saldaña et Rosenblum, 1993). L'effet se produit même si les participants sont prévenus à l'avance de l'existence de cette non-congruence entre les stimuli visuel et auditif et sont invités à répondre uniquement sur la base de ce qu'ils entendent. Toutefois, l'influence de la vision sur la perception de la consonne n'est pas absolue. Tout d'abord, l'effet n'intervient pas toujours pour des stimuli créés de façon appropriée ; on constate une variation importante dans l'intensité des effets obtenus à travers les différentes études réalisées (MacDonald et McGurk, 1978; Massaro, 1987; Green et Kuhl, 1989; Green et Norrix, 1997). Ensuite, certains résultats montrent que des effets contextuels provoqués par des consonnes auditives ne peuvent être reproduits avec un effet McGurk (Roberts et Summerfield, 1981). Enfin, la prévalence de cet effet est diminuée par l'introduction d'un distracteur visuel (Tiippana et Andersen, 2004) ou d'une tâche auditive ou visuelle concurrente (Alsius *et al.*, 2005).

Finalement, même lorsque l'effet McGurk se produit, les percepts qui en résultent ne sont pas considérés comme des parfaits représentants de leur catégorie phonétique. Ainsi, Rosenblum et Saldaña (1992) ont constaté que les sujets trouvaient le /va/ auditif – /va/ visuel plus proche du /va/ purement auditif que le /ba/ auditif – /va/ visuel, bien que les trois stimuli aient chacun été identifiés comme un /va/. Massaro et Ferguson (1993) ont constaté que les sujets, en utilisant une échelle continue de notation, ont évalué le /da/ auditif – /da/ visuel comme plus proche du /da/ auditif que le /ba/ auditif – /da/ visuel. On a constaté de même que les sujets mettaient plus de temps à catégoriser les phonèmes audiovisuels non-congruents que les phonèmes audiovisuels congruents (Massaro *et al.*, 1993; Green et Kuhl, 1991). L'ensemble de ces résultats tend à montrer que même si les sujets ne peuvent identifier un stimulus audiovisuel comme non-congruent, ils ressentent une certaine différence perceptive entre les stimuli congruents et les stimuli non-congruents, bien qu'ils ne parviennent pas à déterminer d'eux-mêmes l'origine de cette différence.

3.5.3 Facteurs linguistiques

Les études visant à étudier l'influence des facteurs linguistiques dans l'intégration de la parole audiovisuelle incongruente font intervenir des mots ou des pseudo-mots auditifs et visuels. Les premières études dans cette voie sont parvenues à reproduire l'effet McGurk pour des mots, avec des effets plus ou moins marqués en fonction des stimuli utilisés (Easton et Basala, 1982; Dekle *et al.*, 1992). Le degré d'influence des indices linguistiques est également variable suivant les études, certaines n'ayant pas trouvé d'influence lexicale ou sémantique (Sams *et al.*, 1998) ; toutefois les études les plus récentes ont démontré une facilitation de l'effet McGurk lorsque la fusion perceptive aboutit à un mot (par comparaison avec un pseudo-mot) et lorsque le mot fusionné concorde avec le contexte sémantique (Brancazio, 2004; Windmann, 2004).

3.6 Mécanismes d'intégration audiovisuelle de la parole

3.6.1 Modèles d'intégration audiovisuelle tardive

On adopte classiquement une distinction entre les modèles d'intégration audiovisuelle *tardive* (ou pré-catégorielle), pour lesquels l'accès au code phonétique se fait dans un premier temps séparément dans chaque modalité, pour être seulement ensuite suivi par une intégration audiovisuelle, et les modèles d'intégration audiovisuelle *précoce* (ou post-catégorielle), pour lesquels l'intégration audiovisuelle de la parole se fait avant l'accès au code phonétique.

Dans la taxonomie des modèles d'intégration audiovisuelle élaborée par Schwartz (2004) sur la base de la typologie originalement proposée par Summerfield (1987), les modèles d'intégration audiovisuelle tardive sont regroupés dans la classe des modèles à *identification séparée* (IS), pour lesquels l'identification phonétique finale est effectuée sur la base des décisions prises séparément dans chacune des modalités.

Les premières études portant sur l'effet McGurk ont proposé un modèle visant à expliquer les percepts audiovisuels incongruents observés (McGurk et MacDonald, 1976; MacDonald et McGurk, 1978). Ce modèle (*visual : place ; auditory : manner*) part du constat que la vision fournit des indices très pertinents pour différencier les lieux d'articulation antérieurs (bilabiales, labiodentales) des lieux d'articulation postérieurs (dentales, palatales, vélaires, uvulaires), tandis que les indices de lieu d'articulation fournis par l'audition sont peu pertinents car facilement dégradables par le bruit. Inversement, l'audition fournit des indices très pertinents pour déterminer le voisement, la nasalité et le mode d'articulation, les indices fournis par la vision étant pratiquement inexistants pour le voisement et la nasalité, et assez restreints pour le mode d'articulation.

Ainsi, dans ce cadre théorique, l'intégration audiovisuelle de la parole se ferait par extraction du lieu d'articulation à partir des informations visuelles, les autres traits articulatoires (voisement, nasalité et mode) étant extraits à partir des informations auditives. Il s'agit donc typiquement d'un modèle d'intégration audiovisuelle de classe IS (intégration tardive), la catégorisation phonétique précédant l'intégration audiovisuelle. Ce modèle permet d'expliquer en grande partie les fusions perceptives observées en cas d'effet McGurk ; cependant il ne fournit pas d'explication satisfaisante pour les combinaisons perceptives (MacDonald et McGurk, 1978).

Un modèle très souvent rencontré dans le domaine de la parole audiovisuelle est le modèle FLMP (*fuzzy logical model of perception*) (Massaro *et al.*, 1993; Massaro, 1987; Massaro et Stork, 1998). Ce modèle effectue une intégration audiovisuelle en deux étapes. La première étape consiste à effectuer une évaluation séparée des signaux auditif et visuel en attribuant à chacun, pour

différents traits phonétiques, une notation sur une échelle continue (il s'agit donc d'un modèle à logique floue). La seconde étape consiste à attribuer à chaque phonème possible une probabilité de perception. Comme chaque phonème est caractérisé par des traits phonétiques parfaitement définis, le calcul de la probabilité de perception des différents phonèmes possibles se déduit directement des notations attribuées aux traits phonétiques lors de la première étape. C'est finalement le phonème auquel aura été attribué la plus grande probabilité de perception qui sera retenu. Bien que la catégorisation effectuée lors de la première étape soit partielle, puisqu'effectuée sur une échelle de notation continue, on s'accorde usuellement pour ranger ce modèle parmi les modèles de classe IS (intégration tardive).

Le FLMP a été appliqué avec succès dans un grand nombre de situations expérimentales, incluant des paradigmes de type McGurk. Toutefois un inconvénient de ce modèle est la présence de paramètres libres, qui doivent par conséquent être évalués à partir des données expérimentales afin de se trouver en adéquation avec ces dernières, ce qui est critiquable sur le principe (Vroomen et de Gelder, 2000; Massaro, 2000). Un inconvénient qui découle directement de la présence de paramètres libres est la bonne adéquation du FLMP avec des données expérimentales qui suggèrent fortement l'intervention de facteurs non phonétiques dans l'intégration audiovisuelle, ces données contredisant par conséquent les hypothèses de base du FLMP (Schwartz, 2004). Afin de résoudre cette contradiction, une évaluation rigoureuse de l'adéquation du FLMP aux données expérimentales a été proposée dans le cadre d'un formalisme probabiliste bayésien (Schwartz, 2004), cette évaluation parvenant à la conclusion que la vraisemblance globale du modèle FLMP est moindre que pour d'autres modèles.

3.6.2 Modèles d'intégration audiovisuelle précoce

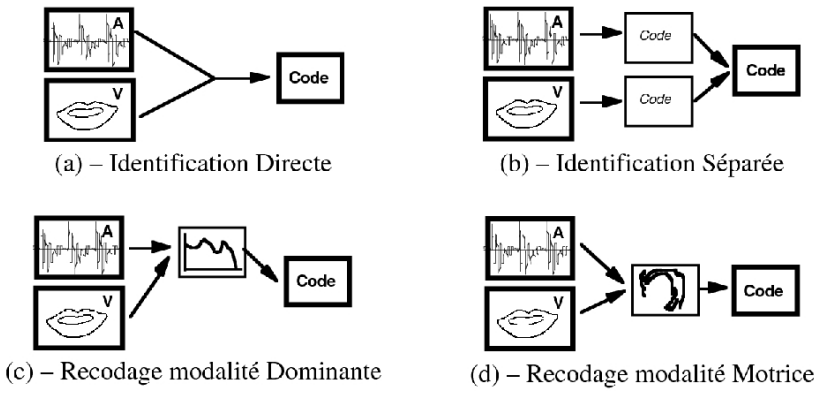


FIG. 3.1 – Modèles possibles de fusion audiovisuelle pour la perception de la parole; d'après Schwartz (2004)

Dans la taxonomie des modèles d'intégration audiovisuelle de la parole proposée par Schwartz (2004), les modèles d'intégration audiovisuelle précoce sont subdivisés en trois classes. Dans la classe des modèles à *identification directe* (ID), l'identification phonétique est effectuée directement à partir des entrées auditive et visuelle, sans traitement préalable des informations sensorielles. Dans ce cadre, un modèle très rudimentaire a été proposé par Summerfield (1987). La méthode consiste à supprimer l'étape de segmentation phonétique et à comparer directement les informations spectro-temporelles auditives et spatio-dynamiques visuelles avec des prototypes de mots, afin d'isoler le mot fournissant la meilleure adéquation avec les informations audiovisuelles.

Ce modèle présente l'avantage d'être facilement utilisable, toutefois son inconvénient majeur est qu'il ne propose aucune explication biologique satisfaisante de l'intégration audiovisuelle.

Les autres classes rassemblent les modèles qui visent à traiter les informations en provenance des deux modalités afin de les recoder dans une modalité commune. Les modèles de *recodage dans la modalité dominante* (RD) se basent sur l'hypothèse que l'audition est la modalité dominante pour la parole, et tentent de traduire les informations visuelles dans une forme de représentation compatible avec les informations auditives (Summerfield, 1987). Cette hypothèse est justifiée par le fait que l'expérience d'une amélioration perceptive audiovisuelle ou d'une perception audiovisuelle illusoire soit ressenti en priorité dans la modalité auditive. Une possibilité de représentation auditive serait dans ce cas l'estimation des fréquences de résonance de l'appareil phonatoire (fréquences formantielles), qui pourrait être effectuée indépendamment à partir des indices auditifs ou visuels.

Enfin, les modèles de *recodage dans la modalité motrice* (RM) viennent s'inscrire dans le cadre de la théorie motrice de la perception de la parole (Liberman *et al.*, 1967; Liberman et Whalen, 2000) en tentant, à partir des informations disponibles dans les modalités auditive et visuelle, de reconstituer les mouvements articulatoires du conduit vocal à l'origine des signaux audiovisuels de parole. Dans le cadre de cette classe de modèles, les caractéristiques principales des gestes articulatoires sont estimées conjointement sur la base des modalités auditive et visuelle, puis une identification phonétique est effectuée à partir de ces données articulatoires.

Plusieurs expériences ont démontré que les indices visuels sont susceptibles d'influencer la catégorisation phonétique auditive, ce qui suggère que l'intégration des informations auditives et visuelles s'effectuerait avant la catégorisation phonétique ; ces études renforcent donc l'hypothèse d'une intégration audiovisuelle de la parole sous-tendue par des mécanismes d'intégration précoce.

Dans ce cadre, l'utilisation d'un continuum auditif entre une syllabe voisée (par exemple /pa/) et une syllabe non-voisée (par exemple /ta/) fournit un paradigme expérimental fin pour tester les propriétés de l'intégration audiovisuelle de la parole. Ainsi ce type de paradigme a permis de démontrer l'influence de la vitesse d'articulation visuelle sur la catégorisation auditive du voisement (Green et Miller, 1985; Brancazio et Miller, 2005). De même, des expériences utilisant l'effet McGurk ont démontré que l'adjonction d'une vélaire visuelle (par exemple /igi/) à un continuum auditif bilabial voisé/non-voisé (par exemple /ibi/-/ipi/) influençait la catégorisation auditive du voisement (Green et Kuhl, 1989; Brancazio *et al.*, 2003).

L'utilisation de la coarticulation est également une façon intéressante d'étudier l'intégration audiovisuelle de la parole ; en effet, dans l'hypothèse où cette intégration serait tardive, c'est-à-dire précédée par une étape de catégorisation phonétique, alors la perception d'un phonème donné ne pourrait pas être influencée par le phonème précédant ou suivant. L'étude de l'effet McGurk sur des consonnes faisant partie de différents contextes vocaliques montre que ce n'est pas le cas, et que la répartition des percepts McGurk obtenus diffère suivant les voyelles qui entourent une consonne donnée (Green *et al.*, 1988; Green et Kuhl, 1991; Sekiyama et Tohkura, 1991; Burnham, 1998; Shigeno, 2000).

3.6.3 Intégration perceptive audio-visuo-motrice de la parole

Un vaste ensemble d'études comportementales et neurofonctionnelles chez l'homme et l'animal converge pour valider l'hypothèse d'une représentation interne phonologique de la parole sous forme de commandes motrices articulatoires (modèles à intégration précoce de type RM), ce qui placerait l'intégration audiovisuelle de la parole dans le cadre de la théorie motrice de la perception de la parole (Liberman *et al.*, 1967; Liberman et Whalen, 2000). Nous développerons

ici les arguments d'ordre comportemental, ceux d'ordre neurofonctionnel étant regroupés dans une partie qui leur est spécialement consacrée (*« Intégration corticale audio-visuo-motrice de la parole »*, § 4.4.3, p. 86).

Les premiers arguments sont de nature développementale. Tout d'abord, pour qu'un enfant parvienne à articuler les sons de parole provenant de son environnement linguistique, il est nécessaire de faire intervenir un mécanisme qui : 1° permette la mémorisation des représentations sensorielles des articulations générées par les autres individus ; 2° permette une comparaison entre les tentatives articulatoires de l'enfant et ces représentations mémorisées ; 3° permette l'utilisation du degré de similarité indiqué par cette comparaison afin de moduler de nouvelles tentatives articulatoires (Hickok et Poeppel, 2004). Par ailleurs de nombreuses expérimentations montrent que ce mécanisme de perception/imitation/production reste actif tout le long de la vie adulte : on observe en effet une dégradation de la production articulatoire après une surdité tardive (Waldstein, 1990), ainsi qu'une modification de la qualité de l'articulation lorsque la rétroaction auditive de la parole est retardée (Yates, 1963) ou dégradée (Houde et Jordan, 1998). De ce point de vue il serait donc cohérent que la représentation phonologique interne se fasse dans un espace de commandes articulatoires, ce qui conditionnerait le codage des percepts audiovisuels de parole à s'exprimer dans ce même espace articulatoire.

Un deuxième faisceau d'arguments est fourni par l'étude de l'effet McGurk. Plusieurs expériences montrent en effet que les lieux d'articulation auditif et visuel sont tous les deux pris en compte lors de l'intégration audiovisuelle de la parole incongruente (McGurk et MacDonald, 1976; Summerfield, 1979; Massaro et Ferguson, 1993). Il apparaît que l'identification audiovisuelle de la parole est basée sur une évaluation systématique des caractéristiques articulatoires les plus susceptibles de correspondre aux informations auditives et visuelles perçues, qu'elles soient congruentes ou pas.

Ainsi, dans le cas particulier où le stimulus incongruent est composé d'un phonème visuel postérieur (dentales, palatales, vélaires ou uvulaires ; par exemple /ka/) et d'un phonème auditif antérieur (bilabiales ou labiodentales ; par exemple /pa/), ce recodage articulatoire est équivalent au modèle « visual : place ; auditory : manner » (McGurk et MacDonald, 1976; MacDonald et McGurk, 1978). Ici, le percept typiquement rapporté est une fusion /ta/, qui est articulatoirement compatible avec la vision d'un phonème postérieur (/ka/) et l'audition d'un phonème non voisé et non nasalisé (/pa/).

Dans le cas inverse d'un phonème visuel antérieur doublé par un phonème auditif postérieur (par exemple /pa/ visuel avec /ta/ auditif), le percept typiquement rapporté est une combinaison /pta/, qui est articulatoirement compatible avec la vision d'un phonème antérieur (/pa/) et l'audition d'un phonème non voisé et non nasalisé (/ta/). On pourrait toutefois s'interroger sur la vraisemblabilité du percept hypothétique /tpa/, qui présente la même compatibilité articulatoire avec les composantes auditive et visuelle. Pour répondre à cette question il faut faire intervenir les contraintes coarticulatoires intrinsèques à l'appareil vocal, qui rendent plus difficile la prononciation d'un /tpa/ que celle d'un /pta/ (Rochet-Capellan et Schwartz, 2007) ; ainsi le percept audiovisuel minimise naturellement la difficulté (et donc la vraisemblabilité) de la production articulatoire.

On peut également se demander pourquoi dans le dernier cas, le phonème auditif /t/ est préservé dans la combinaison audiovisuelle /pta/, tandis que dans le premier cas, le phonème auditif /p/ n'est pas préservé dans la fusion audiovisuelle /ta/ ? Dans l'hypothèse d'une compatibilité articulatoire du percept audiovisuel avec les indices auditifs et visuels, il n'est en réalité pas possible de conserver le phonème /p/ auditif dans le premier cas ; en effet les indices visuels associés indiquent un phonème postérieur (/k/) sans fermeture des lèvres et rendent donc très peu plausible la présence d'un phonème antérieur (comme /p/) avec fermeture des lèvres. Au

contraire, dans le dernier cas, les indices visuels indiquent un phonème antérieur (/p/) avec fermeture des lèvres, ce qui rend possible la présence d'un phonème postérieur /t/ supplémentaire (présent dans la composante auditive), qui serait dans ce cas prononcé après le phonème antérieur /p/, en raison des contraintes coarticulatoires évoquées précédemment (Rochet-Capellan et Schwartz, 2007). Si la présence du /t/ auditif est jugée plausible, le percept audiovisuel résultant sera alors une combinaison /pta/, tandis que si elle jugée non plausible, alors le percept audiovisuel résultant sera à dominance visuelle /pa/. On peut faire l'hypothèse que le degré de pertinence de la présence du /t/ auditif est directement lié à la pertinence des informations auditives, et donc qu'une dégradation de l'audition pourrait éventuellement aboutir à la suppression du /t/ dans le percept audiovisuel (voir notre étude *« Effet McGurk chez le sujet implanté »*, sect. 12.4, p. 271).

3.6.4 Différence d'intégration audiovisuelle entre hommes et femmes

On dispose de peu d'études fournissant une investigation détaillée de la différence d'intégration audiovisuelle de la parole entre les hommes et les femmes. Les paradigmes expérimentaux utilisés dans ce cadre se sont essentiellement focalisés sur la lecture labiale et l'effet McGurk. Ainsi, en fonction des stimuli utilisés et des protocoles expérimentaux, certaines études indiquent que les femmes obtiennent des performances de lecture labiale significativement plus importantes que les hommes (Johnson et Leslie, 1988; Watson *et al.*, 1996), tandis que d'autres ne trouvent pas de différence (Irwin *et al.*, 2006). Par ailleurs, plusieurs études ont démontré une plus grande influence de la composante visuelle de la parole chez les femmes lors de la perception de stimuli de parole audiovisuelle incongruente de type McGurk (Aloufy *et al.*, 1996; Irwin *et al.*, 2006). Cet ensemble de données comportementales est renforcé par des données en neuroimagerie fonctionnelle qui vont dans le même sens et qui fournissent ainsi le point de départ de plusieurs hypothèses permettant d'expliquer ces différences hommes-femmes (voir *« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91).

Bibliographie

Aloufy, S., Lapidot, M. et Myslobodskym, M. (1996). Differences in susceptibility to the "blending illusion" among Native Hebrew and English speakers. *Brain Lang.*, 53(1):51–57.

Alsius, A., Navarra, J., Campbell, R. et Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Curr. Biol.*, 15(9):839–843.

Arnold, P. et Hill, F. (2001). Bisensory augmentation : a speechreading advantage when speech is clearly audible and intact. *Br. J. Psychol.*, 92(Pt 2):339–355.

Aronson, E. et Rosenbloom, S. (1971). Space perception in early infancy : Perception within a common auditory-visual space. *Science*, 172:1161–1163.

Bahrick, L. E., Lickliter, R., Shuman, M., Batista, L. et Grandez, C. (2003). Infant discrimination of voices : Predictions from the intersensory redundancy hypothesis. *In Biennial meeting of the Society for research in child development*, Tampa, FL.

Bahrick, L. E., Netto, D. et Hernandez-Reif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child Dev.*, 69:1263–1275.

Barrera, M. et Maurer, D. (1981). Recognition of mother's photographed face by the three-month-old infant. *Child Dev.*, 52:741–716.

Benoit, C., Mohamadi, T. et Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *J. Speech Hear. Res.*, 37(5):1195–1203.

Binnie, C. A., Montgomery, A. A. et Jackson, P. L. (1974). Auditory and visual contributions to the perception of consonants. *J. Speech Hear. Res.*, 17(4):619–630.

Birnholtz, J. C. et Benacerraf, B. R. (1983). The development of human fetal hearing. *Science*, 222:516–518.

Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *J. Exp. Psychol. Hum. Percept. Perform.*, 30(3):445–463.

Brancazio, L. et Miller, J. L. (2005). Use of visual information in speech perception : evidence for a visual rate effect both with and without a McGurk effect. *Percept. Psychophys.*, 67(5):759–769.

Brancazio, L., Miller, J. L. et Pare, M. A. (2003). Visual influences on the internal structure of phonetic categories. *Percept. Psychophys.*, 65(4):591–601.

Burnham, D. (1998). Language specificity in the development of auditory-visual speech perception. *In* Campbell, R., Dodd, B. et Burnham, D., éditeurs : *Hearing by eye II : Advances in the psychology of speechreading and auditory-visual speech*, pages 27–60. Hove, England : Psychology Press.

Bushnell, I. W. R. (1982). Discrimination of faces by young infants. *J. Exp. Psychol.*, 33:298–308.

Cathiard, M. A. et Tiberghien, G. (1994). Le visage de la parole : une cohérence bimodale temporelle ou configurationnelle. *Psychol. Fr.*, 39(4):357–374.

Conrey, B. et Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *J. Acoust. Soc. Am.*, 119(6):4065–4073.

Cotton, J. C. (1935). Normal "visual hearing". *Science*, 82:592–593.

DeCasper, A. J. et Fifer, W. (1980). Of human bonding : Newborns prefer their mother's voices. *Science*, 208:1174–1176.

DeCasper, A. J. et Prescott, P. A. (1984). Human newborns' perception of male voices : Preference, discrimination and reinforcing value. *Dev. Psychobiol.*, 5:481–491.

DeCasper, A. J. et Spence, M. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.*, 9:133–150.

Dekle, D. J., Fowler, C. A. et Funnell, M. G. (1992). Audiovisual integration in perception of real words. *Percept. Psychophys.*, 51(4):355–362.

Desjardins, R. N., Rogers, J. et Werker, J. F. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *J. Exp. Child Psychol.*, 66(1):85–110.

Dixon, N. F. et Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, 9:719–721.

Dodd, B. (1979). Lipreading in infants : Attention to speech presented in and out of synchrony. *Cognit. Psychol.*, 11:478–484.

Easterbrook, M., Kisilevsky, B., Hains, S. et Muir, D. (1999). Faceness or complexity : Evidence from newborn visual tracking of facelike stimuli. *Infant Behav. Dev.*, 22:17–35.

Easton, R. D. et Basala, M. (1982). Perceptual dominance during lipreading. *Percept. Psychophys.*, 32(6):562–570.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *J. Speech Hear. Res.*, 12(2):423–425.

Erber, N. P. (1975). Auditory-visual perception of speech. *J. Speech Hear. Disord.*, 40(4):481–492.

Fagan, J. F. (1972). Infants' recognition memory for faces. *J. Exp. Child Psychol.*, 14:453–476.

Fant, G. (1973). A note on vocal tract size factors and nonuniform f-pattern scalings. *In* Fant, G., éditeur : *Speech Sounds and Features*, pages 84–93. Cambridge : MIT Press.

Fowler, C. A. et Dekle, D. J. (1991). Listening with eye and hand : cross-modal contributions to speech perception. *J. Exp. Psychol. Hum. Percept. Perform.*, 17(3):816–828.

Grant, K. W. (2001). The effect of speechreading on masked detection thresholds for filtered speech. *J. Acoust. Soc. Am.*, 109(5 Pt 1):2272–2275.

Grant, K. W. et Braida, L. D. (1991). Evaluating the articulation index for audiovisual input. *J. Acoust. Soc. Am.*, 89:2952–2960.

Grant, K. W. et Seitz, P. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *J. Acoust. Soc. Am.*, 108:1197–1208.

Grant, K. W., van Wassenhove, V. et Poeppel, D. (2004). Detection of auditory (crossspectral) and auditory-visual (cross-modal) synchrony. *Speech Comm.*, 44:43–53.

Grant, K. W. et Walden, B. E. (1996). Evaluating the articulation index for auditory-visual consonant recognition. *J. Acoust. Soc. Am.*, 100:2415–2424.

Green, K. P. et Kuhl, P. K. (1989). The role of visual information in the processing of place and manner features in speech perception. *Percept. Psychophys.*, 45(1):34–42.

Green, K. P. et Kuhl, P. K. (1991). Integral processing of visual place and auditory voicing information during phonetic perception. *J. Exp. Psychol. Hum. Percept. Perform.*, 17:278–288.

Green, K. P., Kuhl, P. K. et Meltzoff, A. N. (1988). Factors affecting the integration of auditory and visual information in speech : the effect of vowel environment. *J. Acoust. Soc. Am.*, 84:S155.

Green, K. P. et Miller, J. L. (1985). On the role of visual rate information in phonetic perception. *Percept. Psychophys.*, 38(3):269–276.

Green, K. P. et Norrix, L. W. (1997). Acoustic cues to place of articulation and the McGurk effect : the role of release bursts, aspiration, and formant transitions. *J. Speech Lang. Hear. Res.*, 40(3):646–665.

Hickok, G. et Poeppel, D. (2004). Dorsal and ventral streams : a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99.

Hirsh, I. J. et Sherrick, Jr, C. E. (1961). Perceived order in different sense modalities. *J. Exp. Psychol.*, 62:423–432.

Houde, J. F. et Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354):1213–1216.

Irwin, J. R., Whalen, D. H. et Fowler, C. A. (2006). A sex difference in visual influence on heard speech. *Percept. Psychophys.*, 68(4):582–592.

Johnson, F. M. et Leslie, H. H. (1988). Sex differences in lipreading. *Bull. Psychon. Soc.*, 26(2):106–108.

Jones, J. A. et Jarick, M. (2006). Multisensory integration of speech signals : the relationship between space and time. *Exp. Brain Res.*, 174(3):588–594.

Jones, J. A. et Munhall, K. G. (1997). The effects of separating auditory and visual sources on audiovisual integration of speech. *Can. Acoust.*, 25(4):13–19.

Kim, J. et Davis, C. (2003). Hearing foreign voices : does knowing what is said affect masked visual speech detection ? *Perception*, 32(1):111–120.

Kim, J. et Davis, C. (2004). Investigating the audiovisual speech detection advantage. *Speech Comm.*, 44:19–30.

Kleiner, K. A. (1987). Amplitude and phase spectra as indices of infants' pattern preference. *Infant Behav. Dev.*, 10:49–59.

Kuhl, P. K. et Meltzoff, A. N. (1982). The bimodal development of speech in infancy. *Science*, 218:1138–1141.

Kuhl, P. K. et Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behav. Dev.*, 7:361–381.

Kuhl, P. K. et Meltzoff, A. N. (1988). Speech as an intermodal object of perception. *In* Yonas, A., éditeur : *Perceptual development in infancy : The Minnesota Symposia on child psychology*. Hillsdale, NJ : Erlbaum.

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behav. Dev.*, 17:829–840.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. et Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.*, 74(6):431–461.

Liberman, A. M. et Whalen, D. H. (2000). On the relation of speech to language. *Trends Cogn. Sci.*, 4(5):187–196.

Lisker, L. et Rossi, M. (1992). Auditory and visual cueing of the [+/- rounded] feature of vowels. *Lang. Speech*, 35(Pt 4):391–417.

MacDonald, J. et McGurk, H. (1978). Visual influences on speech perception processes. *Percept. Psychophys.*, 24(3):253–257.

MacKain, K., Studdert-Kennedy, M., Spieker, S. et Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science*, 219:1347–1349.

MacLeod, A. et Summerfield, A. Q. (1987). Quantifying the contribution of vision to speech perception in noise. *Br. J. Audiol.*, 21(2):131–141.

Massaro, D. W. (1987). Speech perception by ear and eye : a paradigm for scientific inquiry. *In* Dodd, B. et Campbell, R., éditeurs : *Hearing by eye : The psychology of lipreading.*, pages 53–83. Hillsdale, NJ : Erlbaum.

Massaro, D. W. (2000). Reply to Vroomen and de Gelder. *Trends Cogn. Sci.*, 4(2):38–39.

Massaro, D. W., Cohen, M. M., Gesi, A., Heredia, R. et Tsuzaki, M. (1993). Bimodal speech perception : An examination across languages. *J. Phon.*, 21:445–478.

Massaro, D. W., Cohen, M. M. et Smeele, P. M. (1996). Perception of asynchronous and conflicting visual and auditory speech. *J. Acoust. Soc. Am.*, 100(3):1777–1786.

Massaro, D. W. et Ferguson, E. L. (1993). Cognitive style and perception : the relationship between category width and speech perception, categorization, and discrimination. *Am. J. Psychol.*, 106(1):25–49.

Massaro, D. W. et Stork, D. G. (1998). Speech recognition and sensory integration. *Am. Scient.*, 86:236.

Maurer, D. et Barrera, M. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Dev.*, 52:196–202.

Maurer, D. et Young, R. E. (1983). Newborns' following of natural and distorted arrangements of facial features. *Infant Behav. Dev.*, 6:127–131.

McGrath, M. et Summerfield, A. Q. (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *J. Acoust. Soc. Am.*, 77(2):678–685.

McGurk, H. et MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. *Infant Behav. Dev.*, 6:313–330.

Miller, C. L., Younger, B. A. et Morse, P. A. (1982). The categorization of male and female voices in infancy. *Infant Behav. Dev.*, 5:143–159.

Miller, G. A. et Nicely, P. (1955). An analysis of perceptual confusions among some english consonants. *J. Acoust. Soc. Am.*, 27(2):338–352.

Mulford, R. (1988). First words of the blind child. *In* Smith, M. et Locke, J., éditeurs : *The emergent lexicon : The child's develoment of a linguistic vocabulary*, pages 293–338. New-York : Academic Press.

Munhall, K. G., Gribble, P., Sacco, L. et Ward, M. (1996). Temporal constraints on the McGurk effect. *Percept. Psychophys.*, 58(3):351–362.

Neely, K. K. (1956). Effect of visual factors on the intelligibility of speech. *J. Acoust. Soc. Am.*, 28:1275–1277.

Pandey, P. C., Kunov, H. et Abel, S. M. (1986). Disruptive effects of auditory signal delay on speech perception with lipreading. *J. Aud. Res.*, 26(1):27–41.

Paré, M., Richler, R. C., ten Hove, M. et Munhall, K. G. (2003). Gaze behavior in audiovisual speech perception : the influence of ocular fixations on the McGurk effect. *Percept. Psychophys.*, 65(4):553–567.

Pascalis, O., de Haan, M., Nelson, C. A. et de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *J. Exp. Psychol. Learn. Mem. Cogn.*, 24:249–260.

Pascalis, O. et de Schonen, S. (1994). Recognition in 3- to 4-day-old human neonates. *Neuroreport*, 5:1721–1724.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C. et Fabre-Genet, M. (1995). Mother's face recognition by neonates : A replication and extension. *Infant Behav. Dev.*, 18:79–85.

Patterson, M. L. et Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behav. Dev.*, 22:237–247.

Patterson, M. L. et Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *J. Exp. Child Psychol.*, 81(1):93–115.

Patterson, M. L. et Werker, J. F. (2003). Two-month old infants match phonetic information in lips and voice. *Dev. Sci.*, 6:191.

Pickens, J., Field, T., Nawrocki, T., Martinez, A., Soutullo, D. et Gonzalez, J. (1994). Full-term and preterm infants' perception of face-voice synchrony. *Infant Behav. Dev.*, 17:447–455.

Querleu, D., Renard, X., Boutteville, C. et Crepin, G. (1989). Hearing by the human fetus ? *Sem. Perinatol.*, 13:409–420.

Reisberg, D., McLean, J. et Goldfield, A. (1987). Easy to hear but hard to understand. *In* Dodd, B. et Campbell, R., éditeurs : *Hearing by eye : The psychology of lipreading.*, pages 97–114. Hillsdale, NJ : Erlbaum.

Robert-Ribes, J., Schwartz, J.-L., Lallouache, T. et Escudier, P. (1998). Complementarity and synergy in bimodal speech : auditory, visual, and audio-visual identification of French oral vowels in noise. *J. Acoust. Soc. Am.*, 103(6):3677–3689.

Roberts, M. et Summerfield, A. Q. (1981). Audiovisual presentation demonstrates that selective adaptation in speech perception is purely auditory. *Percept. Psychophys.*, 30(4):309–314.

Rochet-Capellan, A. et Schwartz, J. L. (2007). An articulatory basis for the labial-to-coronal effect : /pata/ seems a more stable articulatory pattern than /tapa/. *J. Acoust. Soc. Am.*, 121(6):3740–3754.

Rosenblum, L. D. et Saldaña, H. M. (1992). Discrimination tests of visually influenced syllables. *Percept. Psychophys.*, 52(4):461–473.

Rosenblum, L. D., Schmuckler, M. A. et Johnson, J. A. (1997). The McGurk effect in infants. *Percept. Psychophys.*, 59(3):347–357.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. et Foxe, J. J. (2007). Do you see what I am saying ? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb. Cortex*, 17(5):1147–1153.

Saldaña, H. M. et Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Percept. Psychophys.*, 54(3):406–416.

Sams, M., Manninen, P., Surakka, V., Helin, P. et Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences : Effects of word meaning and sentence context. *Speech Comm.*, 26:75–87.

Schwartz, J.-L. (2004). La parole multisensorielle : plaidoyer, problèmes, perspective. *In Journées d'étude sur la parole*. Fès, Maroc.

Schwartz, J.-L., Berthommier, F. et Savariaux, C. (2004). Seeing to hear better : evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2):B69–B78.

Sekiyama, K. et Tohkura, Y. (1991). McGurk effect in non-English listeners : few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *J. Acoust. Soc. Am.*, 90(4 Pt 1):1797–1805.

Sekiyama, K. et Tokhura, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *J. Phon.*, 21:427–444.

Shigeno, S. (2000). Influence of vowel context on the audio-visual speech perception of voiced stop consonants. *Jap. Psychol. Res.*, 42:155–167.

Soken, N. H. et Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Dev.*, 63:787–795.

Spelke, E. S. et Cortelyou, A. (1980). Perceptual aspects of social knowing : Looking and listening in infancy. *In* Lamb, M. et Sherrod, L., éditeurs : *Infant social cognition.*, pages 97–114. Hillsdale, NJ : Erlbaum.

Spelke, E. S. et Owsley, C. (1979). Intermodal exploration and knowledge in infancy. *Infant Behav. Dev.*, 2:13–27.

Sumby, W. H. et Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.*, 26:212–215.

Summerfield, A. Q. (1979). Use of visual information for phonetic perception. *Phonetica*, 36:314–331.

Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. *In* Dodd, B. et Campbell, R., éditeurs : *Hearing by eye : the psychology of lip-reading*, pages 3–52. Erlbaum : Hillsdale, NJ.

Summerfield, A. Q. et Assmann, P. F. (1989). Auditory enhancement and the perception of concurrent vowels. *Percept. Psychophys.*, 45(6):529–536.

Summerfield, A. Q. et McGrath, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *Q. J. Exp. Psychol.*, 36(1):51–74.

Tiippana, K. et Andersen, T. S. (2004). Visual attention modulates audiovisual speech perception. *Eur. J. Cogn. Psychol.*, 16(3):457–472.

van Wassenhove, V., Grant, K. W. et Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, 45(3):598–607.

Vatakis, A. et Spence, C. (2006a). Audiovisual synchrony perception for music, speech, and object actions. *Brain Res.*, 1111(1):134–142.

Vatakis, A. et Spence, C. (2006b). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neurosci. Lett.*, 393(1):40–44.

Vroomen, J. et de Gelder, B. (2000). Crossmodal integration : a good fit is no criterion. *Trends Cogn. Sci.*, 4(2):37–38.

Walden, B. E., Prosek, R. A., Montgomery, A. A., Scherr, C. K. et Jones, C. J. (1977). Effects of training on the visual recognition of consonants. *J. Speech Hear. Res.*, 20(1):130–145.

Waldstein, R. S. (1990). Effects of postlingual deafness on speech production : implications for the role of auditory feedback. *J. Acoust. Soc. Am.*, 88(5):2099–2114.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *J. Exp. Child Psychol.*, 33:514–535.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors : Differentiation of multimodal information. *Psychol. Bull.*, 121:437–456.

Walker-Andrews, A. S. et Lennon, E. (1991). Infants' discrimination of vocal expressions : Contributions of auditory and visual information. *Infant Behav. Dev.*, 14:131–142.

Walton, G. E., Bower, N. J. et Bower, T. G. R. (1992). Recognition of familiar faces by newborns. *Infant Behav. Dev.*, 15:265–269.

Walton, G. E. et Bower, T. G. (1993). Amodal representation of speech in infants. *Infant Behav. Dev.*, 16:233–243.

Watson, C. S., Qiu, W. W., Chamberlain, M. M. et Li, X. (1996). Auditory and visual speech perception : confirmation of a modality-independent source of individual differences in speech recognition. *J. Acoust. Soc. Am.*, 100(2 Pt 1):1153–1162.

Windmann, S. (2004). Effects of sentence context and expectation on the McGurk illusion. *J. Mem. Lang.*, 20:212–230.

Yates, A. J. (1963). Delayed auditory feedback. *Psychol. Bull.*, 60:213–232.

Chapitre 4

Intégration neuronale de la parole et du langage

Sommaire

4.1 Parole et langage 80
4.2 Substrats neuronaux de la parole auditive 81
4.2.1 Reconnaissance de la voix humaine 81
4.2.2 Traitement phonologique auditif de la parole 81
4.3 Substrats neuronaux de la parole visuelle 82
4.3.1 Reconnaissance des visages humains 82
4.3.2 Traitement des mouvements faciaux 83
4.3.3 Traitement phonologique visuel de la parole 84
4.4 Réseau cérébral audio-visuo-moteur de la parole 84
4.4.1 Activations intermodales 84
4.4.2 Intégration des informations auditives et visuelles 85
4.4.3 Intégration corticale audio-visuo-motrice de la parole 86
4.5 Réseau cérébral de traitement du langage 91
4.5.1 Latéralisation hémisphérique du langage 91
4.5.2 Traitement cérébral du langage des signes 93
4.5.3 Traitement du langage dans le cervelet 94
4.5.4 Réseau frontal inférieur droit 95
Bibliographie 97

Introduction

Les données comportementales montrent de façon claire que l'acquisition de la parole et du langage au cours du développement de l'être humain se construit au travers d'une constante alternance entre les fonctions de perception et de production de ces deux composantes fondamentales de la communication parlée. Toutefois la communauté scientifique souffre encore d'un certain manque de recul concernant la compréhension des réseaux corticaux qui sous-tendent la perception de la parole et du langage, dans la mesure

où l'immense majorité des expériences de neuroimagerie destinées à approfondir ce domaine clé ont été réalisées durant ces quinze dernières années, ce qui en fait un champ de connaissance encore jeune.

Quels sont donc les substrats neuronaux qui permettent d'accéder aux représentations phonologiques, lexicales, syntaxiques et sémantiques qui nous sont familières, et ce à partir de l'audition des sons articulatoires produits par un locuteur ou de la vision des mouvements articulatoires de ses lèvres ? Nous nous proposons ici de dresser une vue d'ensemble de l'état actuel des connaissances dans ce domaine.

4.1 Parole et langage

Il nous paraît important de bien expliciter les notions de parole et de langage avant de les aborder sous l'angle des neurosciences. D'après Habib *et al.* (2000), *« pour le neurologue, le psychologue et le linguiste, le langage est un ensemble discret de symboles organisés suivant plusieurs dimensions, dont la combinatoire permet un nombre suffisamment large d'entités pour fournir une étiquette à chacune des notions, concrètes ou abstraites, que l'esprit humain est capable de concevoir (c'est la fonction lexicale du langage), pour organiser ces entités entre elles selon une logique universelle (fonction syntaxique ou grammaire) aboutissant à figurer les liens entre elles, à tisser un réseau de relations entre les mots et les concepts qu'ils représentent (fonction sémantique). »* Ainsi on peut distinguer dans le langage trois niveaux fonctionnels de complexités différentes, qui sont dans l'ordre décroissant la fonction sémantique du langage, la fonction syntaxique puis la fonction lexicale.

La parole peut à son tour être définie comme le support auditif et visuel du langage, c'est-à-dire la combinaison d'informations auditives véhiculées par la voix et d'informations visuelles véhiculées par les mouvements du visage. On peut dès lors faire une distinction entre la parole audiovisuelle, résultant de la combinaison effective de ces deux canaux d'informations, et la parole auditive ainsi que la parole visuelle — parfois désignée de façon restrictive par le terme de « lecture labiale » —, qui n'impliquent qu'un seul canal d'information. Par ailleurs la parole possède une structure phonologique intrinsèque matérialisée par sa décomposition en entités articulatoires distinctes, les phonèmes (voir *« Aspects articulatoires de la phonation »*, sect. 2.2, p. 46 et *« Caractéristiques acoustiques des phonèmes »*, sect. 2.3, p. 52).

En fonction de ces distinctions fonctionnelles, on est amené à se poser les questions suivantes dans le cadre d'une approche neuroscientifique de la perception de la parole et du langage :

- quels sont les réseaux cérébraux qui interviennent dans l'analyse auditive des propriétés spectro-temporelles de la voix et dans l'analyse visuelle des propriétés spatio-dynamiques du visage ?
- comment ces informations sont-elles couplées au niveau cérébral ?
- quels sont les réseaux cérébraux qui permettent de se construire une représentation interne phonologique à partir des informations audiovisuelles véhiculées par la voix et par le visage ?
- sous quelle forme est codée cette représentation interne phonologique ?
- enfin, quels sont les réseaux cérébraux qui sont liés respectivement aux représentations internes lexicale, syntaxique et sémantique du langage ?

Dans la suite de l'exposé, nous tenterons d'apporter des élements de réponse à ces interrogations sous l'éclairage de la littérature scientifique consacrée au sujet.

4.2 Substrats neuronaux de la parole auditive

4.2.1 Reconnaissance de la voix humaine

Belin *et al.* (2000) ont montré qu'il existait des régions du cortex répondant spécifiquement à des stimuli auditifs de voix humaine. Ces régions particulièrement sélectives se situeraient bilatéralement dans la partie antérieure du sillon temporal supérieur (STS). Cette découverte fait apparaître un fort parallèlisme avec le cortex visuel, où des régions sélectives des visages humains ont été observées.

Par ailleurs, ces résultats indiquent une préférence linguistique en faveur de la partie antérieure du STS gauche (Belin *et al.*, 2002), tandis que la partie antérieure du STS droit contribuerait plutôt à l'identification de la voix (Belin et Zatorre, 2003; Fecteau *et al.*, 2004a,b; von Kriegstein et Giraud, 2004). von Kriegstein et Giraud (2004) émettent l'hypothèse que la partie postérieure du STS droit interviendrait plutôt dans l'analyse de la modulation temporelle des voix et des sons. Ainsi, reconnaître la voix en tant que caractéristique spécifique à une personne donnée relèverait plutôt de STS droit antérieur (ventral : "what"), tandis que reconnaître la voix en tant que stimulus acoustique ayant une certaine structure temporelle et spectrale relèverait plutôt du STS droit postérieur (dorsal : "where").

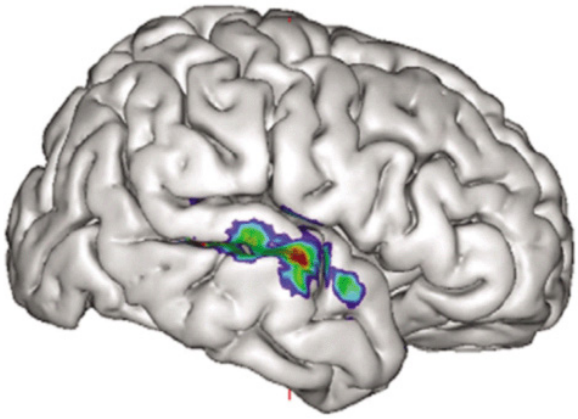


FIG. 4.1 – Activations corticales sélectives aux voix humaines. D'après Belin *et al.* (2004).

4.2.2 Traitement phonologique auditif de la parole

De nombreuses études sur la perception auditive de la parole démontrent l'intervention des aires auditives associatives bilatérales situées dans le sillon temporal supérieur (STS), postérieurement et antérieurement au gyrus de Heschl, s'étendant jusqu'aux extrêmités du STS. Plusieurs d'entre elles tendent à indiquer une dissociation entre les régions moyennes du STS, plutôt impliquées dans l'identification de la parole en tant que signal acoustique, et les régions situés dans la partie plus antérieure du STS, probablement plus impliquées dans la compréhension de la parole, même en cas de dégradation auditive (Belin *et al.*, 2004).

Certains modèles de traitement cortical auditif de la parole ont été proposés (Hickok et Poeppel, 2000, 2004; Skipper *et al.*, 2005; Vigneau *et al.*, 2006). Malgré les différences existant entre ces différents modèles, ceux-ci tendent à s'accorder sur le fait que le traitement phonologique de la parole, c'est-à-dire l'attribution d'une représentation phonétique aux sons de parole, est effectué dans la partie supérieure postérieure du lobe temporal.

4.3 Substrats neuronaux de la parole visuelle

4.3.1 Reconnaissance des visages humains

De nombreuses études ont rapporté l'existence dans le cortex visuel temporal inférieur d'aires sélectives aux visages, situées dans le gyrus fusiforme (Kanwisher *et al.*, 1997; McCarthy *et al.*, 1997; Haxby *et al.*, 2000). Des recherches ultérieures ont permis de préciser le réseau cérébral impliqué dans le traitement visuel des visages humains. Trois régions distinctes émergent de ce réseau cérébral : 1° l'aire occipitale des visages (*occipital face area*), localisée dans le cortex occipital latéral inférieur ; 2° l'aire fusiforme des visages (*fusiform face area*), localisée ventralement sur le gyrus fusiforme ; 3° une région temporale localisée dans la partie postérieure du gyrus temporal supérieur Kanwisher et Yovel (2006).

On peut de plus établir une distinction nette entre les rôles fonctionnels joués par ces trois aires. Ainsi l'aire occipitale des visages serait impliquée dans l'extraction des caractéristiques physiques du visage, sans réponse spécifique à l'identité, avec par ailleurs des réponses similaires entre des visages à l'endroit et des visages à l'envers (Kanwisher et Yovel, 2006). Au contraire, l'aire fusiforme des visages serait spécifiquement impliquée dans la reconnaissance de l'identité du visage (en tant que caractéristique associée à un individu particulier) avec par ailleurs des réponses plus importantes pour des visages à l'endroit que pour des visages à l'envers (Kanwisher et Yovel, 2006). Pour finir, la partie postérieure du gyrus temporal supérieur serait dédiée à des rôles fonctionnels de plus haut niveau comme l'extraction d'indices émotionnels et sociaux à travers l'expression, le regard et la dynamique des traits du visage (Haxby *et al.*, 2000) (voir également ci-après : « *Traitement des mouvements faciaux* », § 4.3.2, p. 83).

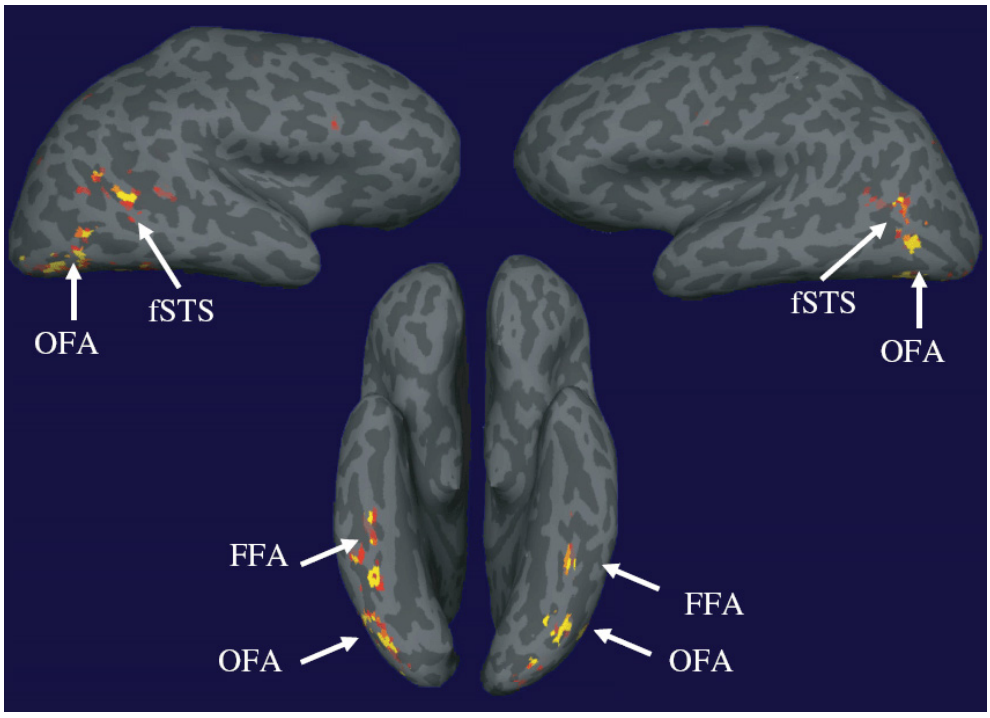


FIG. 4.2 – Activations corticales sélectives aux visages humains. *Notations :* OFA, aire occipitale des visages ; FFA, aire fusiforme des visages ; fSTS, partie postérieure du gyrus temporal supérieur. Adapté d'après Kanwisher et Yovel (2006).

4.3.2 Traitement des mouvements faciaux

Les mouvements du visage sont susceptibles d'émettre de nombreux signaux aux significations sociales diversifiées. Les plus importants de ces signaux sont générés par les mouvements de la bouche et des yeux. Plusieurs études démontrent que l'observation des mouvements des yeux et de la bouche, et plus généralement des mouvements du corps humain, est susceptible d'induire des activations dans de nombreux sites corticaux répartis le long du sillon temporal supérieur (Allison *et al.*, 2000).

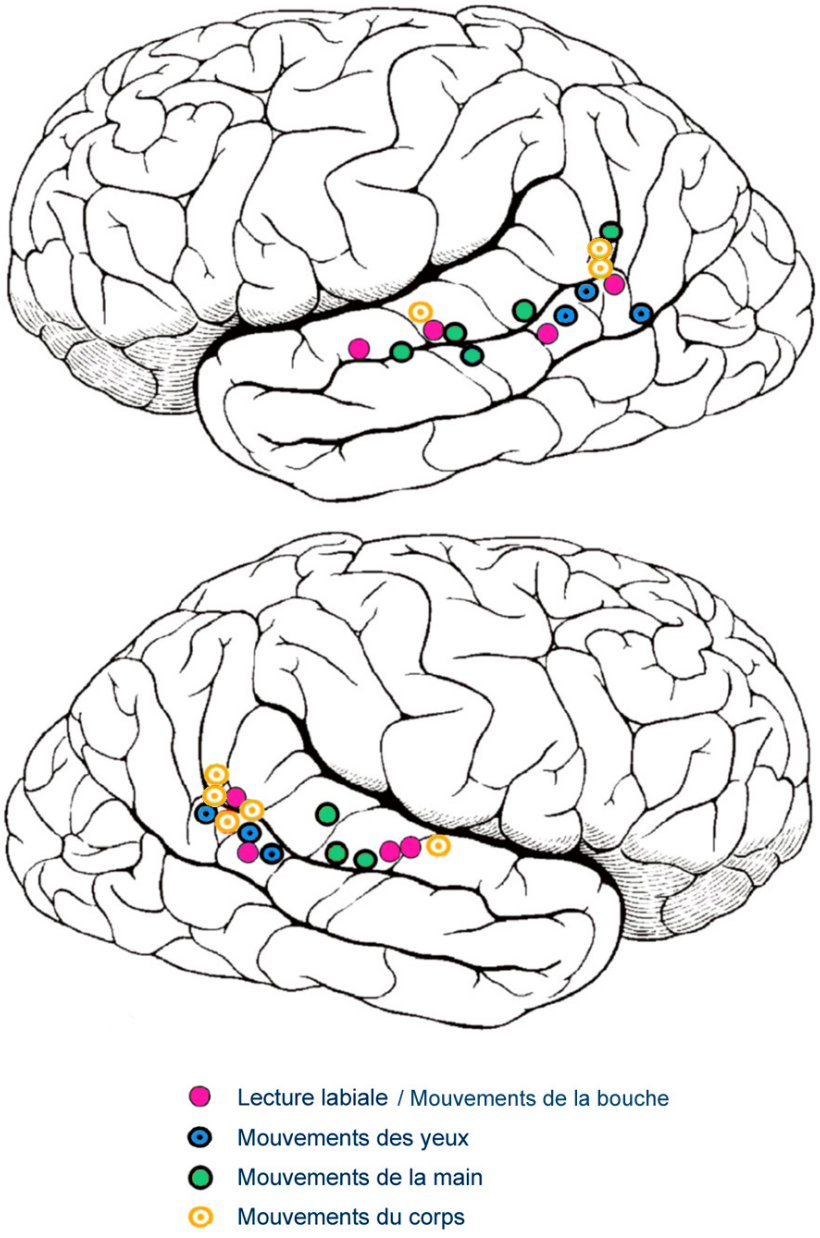


FIG. 4.3 – Activations du sillon temporal supérieur par la vision des mouvements biologiques. Adapté d'après Allison *et al.* (2000).

4.3.3 Traitement phonologique visuel de la parole

Pendant la perception de la parole visuelle, on peut observer une activation de plusieurs régions cérébrales dont en particulier le gyrus temporal supérieur et le gyrus angulaire, ainsi que le gyrus cingulaire postérieur et le cortex frontal médial (Calvert *et al.*, 1997). Ces activations sont à mettre en parallèle avec celles décrites lorsque des sujets observent des mouvements du visage qui pourraient être interprétés comme de la parole, localisées en particulier au niveau du GTS (Puce *et al.*, 1998; Campbell *et al.*, 2001), tandis que l'observation de mouvements de la bouche, linguistiques ou non, induit des activations dans le STS postérieur (Campbell *et al.*, 2001).

Par ailleurs, de nombreuses études montrent que les régions activées par la parole visuelle se superposent en partie à celles activées par la parole auditive, en particulier au niveau des aires auditives secondaires (Sams *et al.*, 1991; Calvert *et al.*, 1997; Callan *et al.*, 2001; Bernstein *et al.*, 2002) et primaires (MacSweeney *et al.*, 2000; Molholm et Foxe, 2005; Pekkola *et al.*, 2005; Ruytjens *et al.*, 2006). Cette superposition suggère l'existence d'afférences multimodales dans ces régions, qui constitueraient dans ce cas un substrat cérébral permettant l'intégration audiovisuelle de la parole (voir *« Intégration des informations auditives et visuelles »*, § 4.4.2, p. 85).

D'après Paulesu *et al.* (2003), qui proposent un modèle anatomique élaboré du réseau cérébral sous-tendant le traitement visuel de la parole, les aires cérébrales impliquées dans la médiation phonologique de la lecture labiale sont disposées selon un réseau temporal et frontal, le sous-réseau temporal faisant intervenir les gyri temporaux supérieur et médian bilatéraux tandis que le sous-réseau frontal fait intervenir le gyrus frontal inférieur gauche.

4.4 Réseau cérébral audio-visuo-moteur de la parole

Les travaux récents en neurophysiologie semblent confirmer les résultats des travaux en psychophysique, qui tendaient à privilégier l'hypothèse du modèle de recodage dans la modalité motrice (cf. *« Intégration perceptive audio-visuo-motrice de la parole »*, § 3.6.3, p. 69) concernant le mécanisme d'intégration audiovisuelle de la parole. Ces études mettent peu à peu en évidence un réseau neuronal permettant de comprendre les actions du locuteur : ce réseau associerait d'une part une région temporale de description multisensorielle des caractéristiques de l'action (au niveau du sillon temporal supérieur STS, point du cortex auditif associatif où convergent les voies auditive et visuelle) et d'autre part une région frontale inférieure permettant de coder et d'interpréter les buts de cette action, et donc les intentions du locuteur (Schwartz, 2004). Cette hypothèse est pour l'instant soutenue par les données expérimentales en imagerie fonctionnelle sur la perception visuelle et audiovisuelle de la parole.

4.4.1 Activations intermodales

On peut observer un fort parallèle entre la vision et l'audition lorsque ces modalités sensorielles sont envisagés sous l'angle de la perception de la parole. Tout d'abord, sur un plan simplement comportemental, la voix joue pour la modalité auditive un rôle similaire à celui du visage pour la modalité visuelle, c'est-à-dire celui d'un canal d'information qui fournit des indices spécifiques concernant l'identité du locuteur et le contenu émotionnel et linguistique de la parole.

Ensuite, sur un plan neurofonctionnel, les cortex visuel et auditif présentent des aires sensorielles spécifiquement activées par respectivement les visages et les voix humaines. Ainsi on trouve des aires sélectives aux visages localisées dans le gyrus fusiforme pour le cortex visuel

(Kanwisher *et al.*, 1997; McCarthy *et al.*, 1997; Haxby *et al.*, 2000), tandis qu'on trouve des aires sélectives aux voix localisées dans le sulcus temporal supérieur antérieur pour le cortex auditif (Belin *et al.*, 2000, 2002).

Ce parallèle est par ailleurs confirmé par les syndromes neuropsychologiques similaires observés en association avec les modalités sensorielles auditive et visuelle. Ainsi certains patients présentant des lésions du lobe temporal peuvent rencontrer des difficultés pour l'identification des visages (prosopagnosie ; Damasio *et al.* (1990); De Renzi *et al.* (1994)) ou des voix (phonagnosie ; Van Lancker *et al.* (1988); Neuner et Schweinberger (2000)).

Des activations intermodales associées à la perception de voix de personnes familières ont été observées dans l'aire fusiforme des visages (von Kriegstein *et al.*, 2005, 2006). Des analyses de connectivité fonctionnelle montrent que l'aire fusiforme des visages est spécifiquement couplée avec le sillon temporal supérieur lors de la reconnaissance auditive des locuteurs familiers. L'activation de l'aire visuelle des visages lors de l'écoute de voix familières pourrait en fait correspondre à un phénomène d'imagerie mentale implicite aboutissant à l'évocation du visage correspondant à la voix.

Inversement, de nombreuses études montrent des activations intermodales du cortex auditif par des stimuli de parole visuelle : des sujets qui regardent des prononciations silencieuses de mots présentent une activation des aires auditives secondaires et même des cortex auditifs primaires (voir références dans *« Traitement phonologique visuel de la parole »*, § 4.3.3, p. 84). Toutefois aucune étude n'a pour l'instant démontré une activation intermodale des aires de la voix du sulcus temporal supérieur antérieur lors d'une présentation de visages articulant silencieusement.

4.4.2 Intégration des informations auditives et visuelles

Lors de la perception de la parole audiovisuelle, on observe une augmentation de l'activité du cortex auditif en comparaison avec la perception de la parole en audition seule (Calvert *et al.*, 1999, 2000).

Les aspects temporels de l'intégration de signaux de parole audiovisuelle ont également été étudiés dans le cas de signaux audiovisuels incongruents de type McGurk, et ce dans plusieurs études neurofonctionnelles incluant des paradigmes en potentiels évoqués (ERP) (Sams *et al.*, 1998; Colin *et al.*, 2002, 2004; Musacchia *et al.*, 2006), en « mismatch-negativity » (MMN) (Sams *et al.*, 1991; Saint-Amour *et al.*, 2007), en imagerie cérébrale par résonance magnétique nucléaire (IRMf) (Jones et Callan, 2003) et en magnétoencéphalographie (MEG) (Möttönen *et al.*, 2002). Ces différentes études concordent pour suggérer une intervention précoce des interactions corticales entre les entrées sensorielles auditive et visuelle, permettant une influence visuelle sur les informations auditives dès le niveau de traitement phonétique (Sams *et al.*, 1998).

Par ailleurs, de nombreuses études anatomiques ont révélé des connexions synaptiques directes entre les cortex visuel et auditif de bas niveau chez le singe. Ainsi les aires visuelles primaires reçoivent des afférences provenant du cortex auditif et de la face supérieure du STS (Falchier *et al.*, 2002; Rockland et Ojima, 2003; Clavagnier *et al.*, 2004) (voir également *« Intégration dans les aires corticales de bas-niveau »*, § 5.4.2, p. 116). De tels résultats confortent l'hypothèse récente selon laquelle il existerait des afférences nerveuses entre les différentes aires corticales sensorielles de bas niveau, autorisant une modulation directe (« feedforward ») du traitement des stimuli sensoriels (Giard et Peronnet, 1999; Schroeder et Foxe, 2002; Foxe *et al.*, 2002; Molholm *et al.*, 2002; Foxe et Schroeder, 2005; Schroeder et Foxe, 2005; Murray *et al.*, 2005). Cet ensemble de résultats expérimentaux suggère fortement l'existence d'une voie corticale permettant l'influence d'informations visuelles sur le traitement des informations auditives,

ce qui rendrait possible des interactions audiovisuelles précoces.

La partie postérieure du sulcus temporal supérieur a été proposée comme un site important pour l'intégration multimodale de la parole audiovisuelle (Calvert *et al.*, 2000; Sekiyama *et al.*, 2003; Wright *et al.*, 2003; Möttönen *et al.*, 2004). Il a également été proposé que l'intégration corticale audiovisuelle pourrait intervenir dans les aires auditives de bas niveau (Möttönen *et al.*, 2004), et même dans des étages sous-corticaux comme le colliculus inférieur (Musacchia *et al.*, 2006; Champoux *et al.*, 2006, 2007).

Latéralisation de l'intégration audiovisuelle de la parole

Les études en neuroimagerie fonctionnelle tendent à mettre en évidence une latéralisation des régions d'intégration audiovisuelle de la parole en faveur de l'hémisphère dominant du langage (Calvert *et al.*, 2000; Calvert, 2001; MacSweeney *et al.*, 2002; Calvert et Lewis, 2004; Callan *et al.*, 2004; Pekkola *et al.*, 2005).

Ces résultats sont confortés par des expériences en neuroimagerie où les activations en conditions de parole audiovisuelle et de parole auditive sont comparées directement. Ainsi une étude récente en TEP a montré des activations du sulcus temporal supérieur fortement latéralisées à gauche en cas de stimulation par de la parole audiovisuelle, tandis que ces activations étaient équilibrées bilatéralement en cas de stimulation par de la parole uniquement auditive (Kang *et al.*, 2006). Une autre étude récente, menée dans un groupe de patients épileptiques ayant subi une implantation de platines d'enregistrement multi-électrodes subdurales au niveau du cortex temporal (incluant le cortex auditif primaire et le gyrus temporal supérieur), a également montré, pour la comparaison entre les conditions audiovisuelle et auditive, des suractivations fortement latéralisées à gauche de la partie postérieure latérale du gyrus temporal supérieur (Reale *et al.*, 2007).

4.4.3 Intégration corticale audio-visuo-motrice de la parole

Réseau cérébral de production/imitation

L'hypothèse d'un processus cérébral qui sous-tendrait à la fois la production et l'imitation d'une action particulière a été particulièrement développée dans le cadre de la théorie des « neurones miroirs ». Les neurones miroirs ont été définis comme une classe particulière de neurones répondant non seulement à l'exécution motrice d'une action mais également à l'observation passive de cette action lorsqu'elle est effectuée par un autre individu. De tels neurones ont été initialement observés chez le singe macaque dans l'aire ventrale prémotrice F5, en réponse à différents types d'actions (attraper, déchirer, tenir) soit effectuées par le singe lui-même, soit effectuées par un expérimentateur situé dans le champ de vision du singe (di Pellegrino *et al.*, 1992; Gallese *et al.*, 1996; Rizzolatti *et al.*, 1996a).

A la suite de cette découverte, il a été suggéré que l'homologue humain de l'aire F5 du singe macaque corresponde aux parties ventrale prémotrice (BA 6) et operculaire (BA 44) de l'aire de Broca, sur la base d'arguments anatomiques, cytoarchitectoniques et neurofonctionnels (Gallese *et al.*, 1996; Rizzolatti et Arbib, 1998). Cette hypothèse s'est révélée cohérente avec la découverte d'un système de neurones miroirs chez l'homme situé dans l'aire de Broca, initialement mis en évidence au travers d'expériences en simulation magnétique transcrânienne (Fadiga *et al.*, 1995), puis par la suite au travers de recherches en tomographie par émission de positons (Rizzolatti *et al.*, 1996b; Grafton *et al.*, 1996) et en imagerie fonctionnelle par résonance magnétique nucléaire (Iacoboni *et al.*, 1999).

D'après Rizzolatti et Arbib (1998), cet ensemble de données indique chez les primates l'existence d'un mécanisme fondamental de reconnaissance des actions effectuées par les autres individus, ce mécanisme étant sous-tendu par la similarité entre les motifs d'activation neuronaux générés dans les aires prémotrices d'une part par l'observation passive d'une action et d'autre part par l'exécution motrice de cette action. C'est ce fort couplage neural entre perception et production des actions qui aurait été à l'origine du développement de la communication chez les animaux, et plus particulièrement du langage chez l'être humain ; par ailleurs ce couplage sensori-moteur plaide fortement en faveur de l'hypothèse d'une représentation interne motrice de la parole (voir ci-après).

Implication sensorielle des aires frontales de la parole

Il est particulièrement important de noter que les aires frontales impliquées dans la motricité de la parole, situées au niveau du cortex frontal moteur et prémoteur — incluant notamment l'aire de Broca —, ont révélé en plus de leur rôle fonctionnel moteur une implication dans le traitement perceptif de la parole. Les premiers indices d'une telle implication proviennent d'études démontrant des activations du système moteur de la parole lors de l'observation passive de divers mouvements de la bouche (Campbell *et al.*, 2001; Nishitani et Hari, 2002; Olson *et al.*, 2002; Callan *et al.*, 2003; Calvert et Campbell, 2003; Paulesu *et al.*, 2003; Buccino *et al.*, 2004; Watkins et Paus, 2004; Pekkola *et al.*, 2006). Des études en stimulation magnétique transcrânienne appliquée au cortex moteur primaire ont également démontré une facilitation de l'excitabilité des msucles articulatoires de la parole lors de la perception de parole visuelle (Sundara *et al.*, 2001; Watkins *et al.*, 2003) ou de parole auditive (Fadiga *et al.*, 2002; Watkins *et al.*, 2003).

Par ailleurs plusieurs études en neuroimagerie fonctionnelle démontrent des activations des aires motrices de la parole aussi bien par la présentation de parole en condition audiovisuelle (Skipper *et al.*, 2005, 2007) qu'en condition auditive seule (Fadiga *et al.*, 2002; Wilson *et al.*, 2004; Pulvermuller *et al.*, 2006; Wilson et Iacoboni, 2006) ou en condition visuelle seule (Campbell *et al.*, 2001; MacSweeney *et al.*, 2000, 2001; Olson *et al.*, 2002; Möttönen *et al.*, 2005). Enfin l'aire de Broca se trouve nettement impliquée lors de tâches de discrimination phonétique de la parole (Démonet *et al.*, 1992; Paulesu *et al.*, 1993, 1996; Démonet *et al.*, 1996; Pugh *et al.*, 1996).

Cet ensemble d'observations plaide donc fortement en faveur d'une mise en jeu d'un réseau « miroir » de production/imitation lors de la perception de la parole (Nishitani et Hari, 2002; Möttönen *et al.*, 2005; Skipper *et al.*, 2007). Ainsi ce réseau permettrait, au cours de la perception audiovisuelle de la parole, la formation d'une représentation interne phonologique de la parole sous forme de commandes motrices articulatoires. L'existence de ce système de production/imitation est par ailleurs en accord avec les hypothèses proposées dans le cadre de la théorie motrice de la perception de la parole (Liberman *et al.*, 1967; Liberman et Mattingly, 1985) (voir *« Intégration perceptive audio-visuo-motrice de la parole »*, § 3.6.3, p. 69).

Implication motrice des aires temporales de la parole

Dès le XIX^{e} siècle, Wernicke a remarqué que les patients présentant des lésions du cortex temporal postérieur rencontraient non seulement des difficultés de compréhension de la parole, mais également des déficits de production (Wernicke, 1874). Sur la base de ces observations, il a fait l'hypothèse que le gyrus temporal supérieur postérieur jouait un rôle à la fois dans la perception et la production de la parole.

Les données lésionnelles modernes supportent l'hypothèse initiale de Wernicke supposant l'implication des aires temporales postérieures dans la production de la parole. Par exemple, les aphasiques de Wernicke présentent des lésions temporales postérieures relativement larges et

montrent typiquement une production de parole fluide mais paraphasique (en plus de leurs difficultés de compréhension) (Damasio, 1991; Goodglass, 1993), ce qui supporte l'argument général qu'une atteinte des aires temporales postérieures (classiquement considérées comme sensorielles) puisse affecter la production de la parole. De façon intéressante, lorsque la lésion est restreinte au gyrus temporal supérieur gauche (aire classique de Wernicke), le déficit constaté concerne la production mais pas la compréhension de la parole (Damasio et Damasio, 1980) ; de telles lésions sont associées avec une aphasie de conduction, un syndrome caractérisé par de fréquentes paraphasies phonémiques dans la production de la parole en parallèle avec une bonne compréhension (Damasio et Damasio, 1980; Goodglass, 1992, 1993; Hickok et Poeppel, 2000). Plusieurs auteurs ont argumenté que les aphasiques de conduction présentent un déficit impliquant l'étape d'encodage phonologique de la production de la parole (Wilshire et McCarthy, 1996). Cela suggère l'implication du gyrus temporal supérieur dans le traitement phonologique lors de la production de la parole (Hickok et Poeppel, 2000; Hickok, 2001; Hickok et Poeppel, 2004).

Ces données neuropsychologiques sont corroborées par les données issues des expériences en neuroimagerie fonctionnelle, ce qui donne un poids important à l'hypothèse de l'implication des régions temporales postérieures dans la production de la parole. Ainsi l'activation du gyrus et du sulcus temporal supérieur postérieur lors de la production articulatoire ou silencieuse de la parole a été mise en évidence par des études en magnétoencéphalographie (Salmelin *et al.*, 1994; Levelt *et al.*, 1998), en tomographie par émission de positions (Wise *et al.*, 1999; Indefrey et Levelt, 2000; Wise *et al.*, 2001) et en imagerie fonctionnelle par résonance magnétique nucléaire (Hickok et Poeppel, 2000; Okada *et al.*, 2003; Okada et Hickok, 2006). Les activations trouvées ne sont pas provoquées par les entrées sensorielles auditives dans la mesure où des tâches similaires supprimant tout retour auditif ont démontré les mêmes activations. Par exemple, Hickok et Poeppel (2000) et Okada *et al.* (2003) ont utilisé un paradigme de production silencieuse (mentale) de dénomination d'objets. En outre, certaines de ces études ont permis de localiser avec précision une région dans la partie dorsale du gyrus temporal supérieur postérieur gauche, située à la jonction des gyri temporal supérieur et supramarginal, qui présente des activations lors des mouvements articulatoires de la parole (Wise *et al.*, 1999, 2001).

Les données obtenues en neuroimagerie fonctionnelle tendent de plus en plus à supporter l'idée d'une implication du gyrus temporal supérieur postérieur gauche dans le traitement phonologique pendant la production de la parole. D'après Wise *et al.* (1999), cette région contribuerait à l'interfaçage des représentations auditives de la parole avec leurs contreparties motrices, les informations de parole codées par le système auditif étant finalement transcodées en actes moteurs articulatoires. Ce processus de transcodage nécessiterait alors une maintenance des représentations de parole grâce à un système de mémoire de travail phonologique. Cette proposition est renforcée par les études de Paulesu *et al.* (1993) et Démonet *et al.* (1994), qui au travers de protocoles destinés à investiguer les substrats neuronaux de la mémoire de travail phonologique, ont montré des activations spécifiques au niveau de la jonction entre le cortex temporal supérieur postérieur et le gyrus supramarginal inférieur. Ces auteurs proposent que cette région constitue le substrat neuronal du stockage transitoire des représentations phonologiques.

Toutes les régions temporales mentionnées ci-avant sont plus classiquement considérées comme des aires crucialement impliquées dans la perception de la parole (Zatorre *et al.*, 1996; Binder *et al.*, 2000; Scott *et al.*, 2000). On peut donc se demander si ce sont les mêmes régions du gyrus temporal supérieur qui sont impliquées aussi bien dans la perception que dans la production de la parole. De nombreuses études apportent une réponse à cette interrogation dans la mesure où elles confirment bel et bien la co-activation des parties postérieure et dorsale du gyrus temporal supérieur gauche à la fois lors de la production et de la perception de la parole, au travers de

tâches variées comme l'écoute d'un récit et la génération de verbes (Papathanassiou *et al.*, 2000), la perception et la répétition silencieuse de pseudo-mots (Buchsbaum *et al.*, 2001), la perception et la répétition de phrases (Hickok *et al.*, 2003), ainsi que la perception et la répétition de mots simples (Okada et Hickok, 2006).

Réseau cérébral d'intégration audio-visuo-motrice de la parole

D'après l'exposé des deux paragraphes précédents, on dispose d'un large ensemble d'observations expérimentales qui montrent : 1° l'implication fonctionnelle des aires frontales inférieures, dont le rôle est classiquement considéré comme moteur, dans des processus de perception audiovisuelle de la parole (y compris du traitement phonologique de la parole) ; 2° l'implication fonctionnelle des aires temporales supérieures postérieures, dont le rôle est classiquement attribué à la perception audiovisuelle de la parole, dans des processus d'articulation motrice de la parole.

Ces observations soulèvent fortement l'idée d'une coopération fonctionnelle étroite entre les réseaux de perception et de production de la parole ; cette idée est par ailleurs renforcée par la mise en évidence de connexions réciproques entre les régions temporales supérieures postérieures et les régions prémotrices inférieures, via le faisceau arqué (arcuate fasciculus) (Binder et Frost, 1998 ; Belin et Zatorre, 2000 ; Hickok et Poeppel, 2000). Pour finir, cette idée est particulièrement cohérente avec la découverte des systèmes de « neurones miroirs » sous-tendant à la fois la perception et la reproduction des actions effectuées par les autres, ainsi qu'avec l'hypothèse d'une représentation interne de la parole sous forme de commandes motrices articulatoires (Liberman *et al.*, 1967 ; Liberman et Whalen, 2000) (voir *« Modèles d'intégration audiovisuelle précoce »*, § 3.6.2, p. 68).

D'après une méta-analyse récente effectuée par Vigneau *et al.* (2006) sur les activations phonologiques et linguistiques de l'hémisphère gauche observées dans 129 études neurofonctionnelles, les réseaux cérébraux impliqués dans le traitement phonologique de la parole peuvent être répartis en deux composantes : 1° un réseau fronto-temporal faisant intervenir une boucle audio-motrice de traitement phonologique de la parole, ce réseau associant d'une part un sous-réseau temporal (perceptif) composé du gyrus de Heschl, du planum temporale et de la partie postérieure du gyrus temporal supérieur, et d'autre part un sous-réseau frontal (moteur) composé de l'aire motrice de la bouche et du cortex prémoteur inférieur ; 2° un réseau fronto-pariétal dédié à la memoire de travail phonologique.

Cette décomposition vient compléter le modèle neurolinguistique proposé par Hickok et Poeppel (2004), qui distinguent : 1° une voie auditive ventrale reliant les aires temporales supérieures et inférieures, cette voie étant dédiée à l'attribution d'une dimension sémantique aux sons ; 2° une voie auditive dorsale reliant les aires temporales supérieures d'abord aux jonctions temporo-pariéto-occipitales puis aux régions frontales prémotrices, cette voie étant dédiée à l'attribution d'une représentation motrice articulatoire aux sons.

Par ailleurs, Skipper *et al.* (2005) proposent explicitement un modèle d'intégration audio-visuo-motrice de la parole, en indiquant un réseau cortical divisé en un sous-réseau temporal supérieur postérieur et un sous-réseau frontal, ce dernier étant composé de la partie supérieure de la pars opercularis de l'aire de Broca, des cortex prémoteurs dorsal et ventral supérieur ainsi que du cortex moteur adjacent.

Ainsi, pour esquisser une synthèse des modèles proposés et des observations rapportées dans la littérature, le traitement phonologique de la parole serait sous-tendu par un réseau cérébral fortement latéralisé dans l'hémisphère dominant du langage. Ce réseau serait subdivisé en 1° une

composante temporale focalisée au niveau de la partie postérieure du gyrus temporal supérieur ; 2° une composante frontale focalisée au niveau de la partie prémotrice des gyri frontaux inférieur et moyen (incluant la partie postérieure de l'aire de Broca) ; 3° un relais temporo-frontal constitué du faisceau arqué et d'une région située entre la partie ventrale du gyrus supramarginal et la partie dorsale du gyrus temporal supérieur postérieur.

Le gyrus temporal supérieur postérieur serait en priorité un relais sensoriel d'intégration multimodale des informations auditives et visuelles associées à la parole, tandis que les aires frontales prémotrices sous-tendraient en priorité les représentations motrices articulatoires inférées à partir de la perception audiovisuelle de la parole. Par ailleurs ces relais respectivement sensoriels (temporal) et moteurs (frontal) se trouveraient en étroite interaction lors de la perception de la parole, que celle-ci soit audiovisuelle ou simplement auditive ou visuelle. Cette interaction serait sous-tendue par un substrat neuronal autorisant le stockage transitoire des représentations phonologiques internes de la parole, ce substrat étant probablement localisé au niveau de la jonction entre les gyri temporal supérieur postérieur et supramarginal. C'est cette étroite interaction corticale entre les centres de traitement audiovisuel et moteur qui permettrait d'établir une relation d'équivalence entre la structure audiovisuelle des éléments de parole perçus et les mouvements articulatoires susceptibles de reproduire cette structure audiovisuelle.

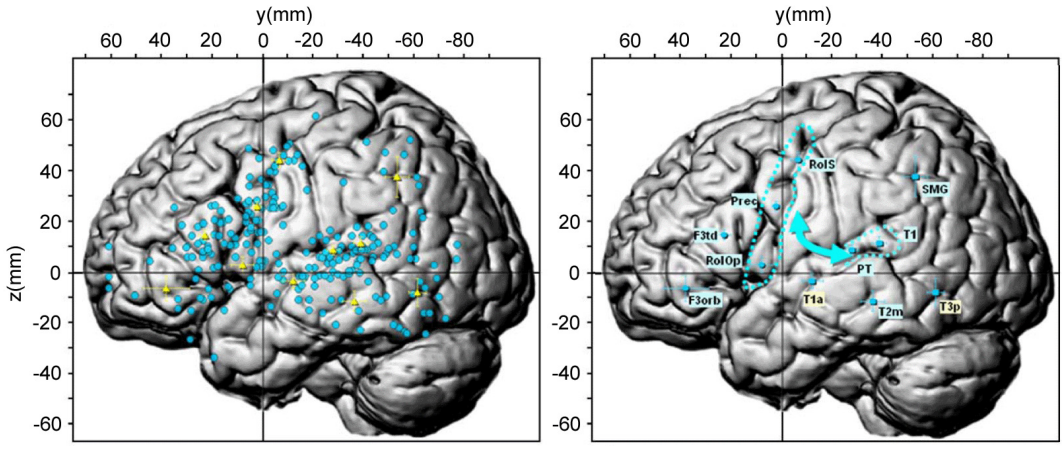


FIG. 4.4 – Réseau cortical de traitement phonologique de la parole. Méta-analyse de 247 pics d'activation corticale répartis dans 86 contrastes provenant de 45 études neurofonctionnelles portant sur le traitement phonologique de la parole. *A gauche* : projection sagittale des pics d'activation ; *à droite* : illustration de la boucle audio-motrice temporo-frontale. *Notations* : RolS, scissure rolandique ; RolOp, opercule rolandique ; F3td/F3orb, pars triangularis dorsale/pars orbicularis du gyrus frontal inférieur ; Prec, gyrus precentral ; SMG, gyrus supramarginal ; T1/T1a, gyrus temporal supérieur postérieur/antérieur ; T2m, gyrus temporal moyen médian ; T3p, gyrus temporal inférieur postérieur ; PT, planum temporale. Adapté d'après Vigneau *et al.* (2006).

4.5 Réseau cérébral de traitement du langage

4.5.1 Latéralisation hémisphérique du langage

Latéralisation du langage verbal

Les premières observations faites par Broca et Wernicke ont suggéré une latéralisation des fonctions linguistiques dans l'hémisphère cérébral gauche (Broca, 1861; Wernicke, 1874). Ces observations ont été confirmées par la suite par l'utilisation du test de Wada. Le principe de ce test est le suivant : afin d'établir un diagnostic de latéralisation hémisphérique du langage, par exemple avant une intervention neurochirurgicale, on injecte de l'amytal de sodium (anesthésique à effet bref) dans une artère carotide, ce qui a pour effet d'inactiver l'hémisphère cérébral placé du côté de l'artère pendant quelques secondes. On peut alors demander au sujet d'effectuer une tâche de traitement linguistique (par exemple une dénomination d'images), puis après disparition de l'effet anesthésique, on répète l'opération de l'autre côté. Dans l'ensemble, les observations réalisées dans le cadre du test de Wada concordent parfaitement pour indiquer une très nette latéralisation du langage dans l'hémisphère gauche chez les sujets droitiers (Rasmussen et Milner, 1977; Rutten *et al.*, 2002; Sabbah *et al.*, 2003; Hirata *et al.*, 2004). Toutefois cette dominance est moins nettement marquée dans le cas des sujets gauchers, une certaine proportion de ces sujets — variable suivant les études — montrant plutôt des activations bilatérales voire une latéralisation dans l'hémisphère droit.

La latéralisation du langage dans l'hémisphère gauche a également été confirmée par l'utilisation des techniques modernes d'imagerie fonctionnelle (TEP, IRMf et MEG), qui ont montré des activations plus importantes dans l'hémisphère gauche lors de diverses tâches spécifiques du langage. Ces asymétries sont par ailleurs très marquées chez les sujets droitiers mais légèrement moins marquées chez les sujets gauchers (Démonet *et al.*, 1992; Klein *et al.*, 1995; Zatorre *et al.*, 1996; Frost *et al.*, 1999; Rutten *et al.*, 2002; Sabbah *et al.*, 2003; Hirata *et al.*, 2004; Kirveskari *et al.*, 2006). Les activations corticales liées au langage sont toutefois rarement restreintes à l'hémisphère gauche ; en réalité un certain nombre d'études réalisées chez des sujets droitiers ont montré des activations bilatérales sans asymétrie (Démonet *et al.*, 1994; Grasby *et al.*, 1994).

Plusieurs hypothèses sont avancées pour expliquer la différence observée entre les deux hémisphères cérébraux. Price *et al.* (1999) proposent que les expériences en neuroimagerie fonctionnelle réalisées chez des sujets sains droitiers montrent un ensemble de zones corticales dont l'activation n'est pas toujours indispensable à la réalisation des tâches demandées. Ainsi seules les études lésionnelles offriraient la possibilité de détecter les aires réellement indispensables, très probablement situées dans l'hémisphère gauche.

Kimura (1973) a émis l'hypothèse que le biais à droite pour la latéralisation manuelle préférentielle observée dans la population humaine serait corrélé à l'efficacité des aptitudes spatiales et temporelles dans l'aire prémotrice gauche, ce qui expliquerait la contribution préférentielle de l'hémisphère gauche pour les compétences linguistiques. Cette hypothèse peut être mise en parallèle avec l'hypothèse « temporelle », qui attribuerait à l'hémisphère gauche une spécialisation dans le traitement des signaux acoustiques à fluctuations rapides (Tallal, 1980; Tallal *et al.*, 1993; Eden *et al.*, 1996; Poldrack *et al.*, 1999; Zatorre *et al.*, 2002; Démonet *et al.*, 2005). Cette spécialisation temporelle de l'hémisphère gauche serait complémentaire d'une spécialisation de l'hémisphère droit dans l'analyse spectrale fine des signaux sonores (Zatorre *et al.*, 2002).

Une seconde hypothèse, tout à fait compatible avec la précédente, est l'hypothèse « catégorielle », selon laquelle l'hémisphère gauche serait spécialisé dans le stockage sur le long terme des représentations phonétiques, ce qui permettrait la catégorisation des phonèmes par l'inter-

médiaire des hauts niveaux de traitement cognitif (mécanismes « top-down ») (Kosslyn, 1987; Celsis *et al.*, 1999; Parrot *et al.*, 1999; Ruff *et al.*, 2002). Cette compétence perceptive spécifique serait complémentaire à la coordination, qui consiste à évaluer la distance perceptive entre deux phonèmes (Kosslyn, 1987).

Latéralisation du langage signé

La latéralisation du traitement du langage s'applique non seulement au langage verbal oral et écrit, mais également au langage signé utilisé pour communiquer avec les personnes souffrant d'une surdité. L'organisation cérébrale des aires permettant de traiter le langage signé présente en effet plusieurs similarités importantes avec le langage verbal (voir *« Traitement cérébral du langage des signes »*, § 4.5.2, p. 93), en particulier concernant la latéralisation hémisphérique des aires concernées. Cette latéralisation est spécifique du traitement linguistique ; en effet, chez les patients sourds utilisant le langage des signes, le traitement est latéralisé à gauche pour des signes ayant une signification linguistique tandis qu'il devient latéralisé à droite pour des signes sans signification linguistique (Bellugi *et al.*, 1988). Ces observations sont également renforcées par des données lésionnelles chez des signeurs atteints de surdité. Chez ces personnes, des lésions de l'hémisphère gauche provoquent une aphasie de la langue des signes (Poizner *et al.*, 1990), tandis que les signes non linguistiques sont toujours compris correctement (Bellugi *et al.*, 1983). Toutefois le débat reste ouvert sur l'hypothèse d'une corrélation entre le degré de latéralisation du traitement du langage des signes et la latéralisation du traitement du langage parlé (Hickok *et al.*, 1998; Neville *et al.*, 1998; Paulesu et Mehler, 1998; Rönnberg *et al.*, 2000; Corina et McBurney, 2001; MacSweeney *et al.*, 2002).

Indices phylogénétiques d'une latéralisation du langage

Plusieurs auteurs ont proposé l'idée que les vocalisations des singes pourraient constituer des précurseurs de la parole humaine, en partie en raison du fait qu'elles fournissent des informations cruciales pour les autres membres de l'espèce (Cheney et Seyfarth, 1990; Prell *et al.*, 2002). Ces vocalisations sont en effet utilisées dans le cadre d'interactions sociales, de comportements de recherche de nourriture, et de comportements de défense ou de fuite à l'approche d'un prédateur.

Dans ce cadre théorique, plusieurs études comportementales ont démontré que les singes présentaient un comportement d'orientation auditive spécifique du son entendu. Lorsque des vocalisations spécifiques de l'espèce sont présentées derrière le singe, celui-ci tourne la tête de façon préférentielle du côté droit, afin d'orienter l'oreille droite vers la source sonore. Ce comportement est spécifique des vocalisations de l'espèce, une orientation préférentielle de l'oreille gauche étant observée pour des sons contrôles comme des vocalisations inversées temporellement ou des chants d'oiseaux ; par ailleurs il n'est observé que chez les sujets adultes (Hauser et Andersson, 1994; Ghazanfar *et al.*, 2001).

Ces observations comportementales seraient révélatrices d'un traitement des signaux de communication propres à l'espèce effectué de façon préférentielle dans l'hémisphère gauche (correspondant à l'oreille droite). Plusieurs études lésionelles chez le macaque sont venues confirmer cette hypothèse, en montrant que certaines parties du gyrus temporal supérieur gauche sont cruciales pour la perception des vocalisations spécifiques de l'espèce, cette spécialisation particulière n'étant pas observée dans l'hémisphère droit (Heffner et Heffner, 1984; Poremba *et al.*, 2004).

Ainsi, dans l'hypothèse où les vocalisations des singes seraient effectivement un précurseur phylogénétique de la parole humaine, cet ensemble d'observations comportementales et neurophysiologiques vient renforcer l'idée d'une latéralisation marquée à gauche chez les primates humains et non-humains pour le traitement des signaux permettant la communication interindividuelle.

Différences de latéralisation entre hommes et femmes

De nombreuses études portant sur le réseau cérébral fonctionnellement impliqué dans la perception de la parole et du langage tendent à indiquer des différences entre les motifs d'activation observés chez les hommes et les femmes, ces études venant renforcer les données comportementales disponibles sur le sujet (voir *« Différence d'intégration audiovisuelle entre hommes et femmes »*, § 3.6.4, p. 71). Bien que ces motifs diffèrent en fonction du protocole expérimental utilisé, les hommes présentent usuellement des latéralisations plus marquées dans l'hémisphère gauche, tandis que les activations chez les femmes sont plus diffuses et symétriques (Shaywitz *et al.*, 1995; Jaeger *et al.*, 1998; Kansaku *et al.*, 2000; Kansaku et Kitazawa, 2001; Phillips *et al.*, 2001; Rossell *et al.*, 2002).

En se plaçant dans le cadre du modèle d'activation-vérification de traitement du langage proposé par Strelnikov (2007), cette différence pourrait indiquer que les femmes utilisent de façon plus marquée des stratégies intégratives et prédictives que les hommes. Ainsi, plusieurs études portant sur les aspects émotionnels du traitement de la parole ont démontré des activations plus rapides et plus étendues chez les femmes que chez les hommes (Schirmer *et al.*, 2002, 2005). Cette différence pourrait également être révélatrice d'une plus forte implication de l'hémisphère droit chez les femmes lors du traitement cortical de la parole et du langage.

Jaeger *et al.* (1998) ont fait l'hypothèse que les différences hommes-femmes dans le traitement cérébral du langage s'accentuent lorsque la complexité linguistique de la tâche augmente. Cette hypothèse peut être replacée dans le cadre du modèle d'activation-vérification du traitement du langage, qui indique que ce sont les stimuli les plus proches d'un flux de parole continue qui impliquent le plus des stratégies cognitives de type prédictif et intégratif (Strelnikov, 2007). Par conséquent, on pourrait faire l'hypothèse d'une plus grande supériorité des femmes lorsque les stimuli atteignent un niveau de complexité similaire à celui de la parole continue.

4.5.2 Traitement cérébral du langage des signes

Plusieurs données lésionelles chez des personnes sourdes rapportent des aphasies dans l'utilisation du langage des signes (Poizner *et al.*, 1987; Corina, 1998). Ces aphasies présentent par ailleurs des caractéristiques très similaires aux aphasies verbales documentées chez les personnes normo-entendantes, des lésions de l'hémisphère gauche dans les aires de Broca et de Wernicke provoquant sur le traitement du langage des signes chez les personnes sourdes des effets comparables à ceux qu'elles auraient eu sur le traitement de la parole chez des personnes normo-entendantes (Hickok *et al.*, 1996; Campbell et MacSweeney, 2004).

Ces similitudes entre les substrats neuronaux du langage des signes et du langage parlé ont par la suite été renforcées par des études en neuroimagerie fonctionnelle. Ainsi l'aire de Broca, normalement impliquée dans la production du langage parlé, est impliquée dans la production du langage des signes chez les personnes sourdes (McGuire *et al.*, 1997; Corina et McBurney, 2001). De même, chez les personnes sourdes, l'aire de Wernicke est spécifiquement activée lors de la perception de signes ayant une signification (Söderfeldt *et al.*, 1994, 1997; Neville *et al.*, 1998; Nishimura *et al.*, 1999; Petitto *et al.*, 2000).

De façon générale, il y a consensus sur le fait que le traitement du langage des signes active des régions normalement impliquées dans le traitement du langage parlé (incluant les aires de Broca et de Wernicke), mais également sur le fait que certaines régions ne servant pas au traitement du langage parlé sont spécifiquement recrutées lors du traitement du langage des signes. C'est en particulier le cas lors du traitement de phrases signées, une tâche riche en informations spatiales, où l'on observe des activations des lobes pariétaux (MacSweeney *et al.*, 2002).

4.5.3 Traitement du langage dans le cervelet

L'éventail des rôles fonctionnels du cervelet s'élargit à mesure que les recherches en neuroimagerie progressent. A la fin du XIX^{e} siècle il était établi de façon certaine que le cervelet jouait un rôle important dans les fonctions motrices du corps humain. D'après Bloedel et Bracha (1997), les conceptions scientifiques concernant le rôle du cervelet ont évolué suivant cinq phases principales durant le XX^{e} siècle : 1^{o} coordination des mouvements volontaires orientés vers un but précis et orientation du corps et de la tête dans l'espace ; 2^{o} régulation et intégration des informations sensorielles pour l'organisation des réflexes proprioceptifs et cutanéomusculaires ; 3^{o} régulation des mouvements vestibulo-oculaires et posture de la tête ; 4^{o} apprentissage de réponses aversives conditionnées ; 5^{o} modifications comportementales aux niveaux linguistique, cognitif et affectif.

Entre les années 1970 et 1980, les nouvelles techniques de neuroimagerie ont suscité un intérêt croissant pour les implications fonctionnelles non-motrices du cervelet et plusieurs études de cas ont indiqué le possible rôle pathogène du cervelet dans des dysfonctionnements cognitifs variés (Boltshauser et Isler, 1977; Hamilton *et al.*, 1983; Fehrenbach *et al.*, 1984). Des études ultérieures ont en outre montré que même en l'absence de toute activité motrice, on pouvait observer des activations du cervelet lors de tâches cognitives et linguistiques variées (Petersen *et al.*, 1988). Ainsi les avancées des recherches en neurosciences de ces trente dernières années ont peu à peu fait émerger le concept de « cognition cérebelleuse » (Marien *et al.*, 2001a).

Connectivité cérébello-corticale

Durant l'évolution humaine, le cervelet a subi des changements structurels importants, dont en particulier un élargissement du noyau denté. Ces modifications structurelles ont été accompagnées par l'apparition de nouvelles connexions cérebello-corticales et sont probablement à l'origine, comme postulé par Leiner *et al.* (1986, 1989), de l'évolution fonctionnelle du cervelet chez l'homme. Ces nouvelles connexions relient entre elles certaines des parties phylogénétiquement nouvelles du cerveau et du cervelet, à savoir le cervelet latéral et le cortex préfrontal. En particulier, Engelborghs *et al.* (1998) ont montré que le cervelet latéral envoie des projections au cortex préfrontal controlatéral dans les aires de Brodmann 6, 44 et 45 (ce qui inclut l'aire de Broca et l'aire motrice supplémentaire) via des relais thalamiques (noyaux ventraux antérieurs et intermédiaires). Leiner *et al.* (1986, 1989) ont mis en évidence les rétroprojections correspondantes, passant par le pont et par la partie parvocellulaire du noyau rouge. Cette découverte de voies neurales réciproques entre le cervelet et les aires frontales du langage représente un saut conceptuel important dans la contribution du cervelet aux processus linguistiques.

Fonctions linguistiques du cervelet

Les études récentes sur le rôle fonctionnel du cervelet dans les processus linguistiques mettent en évidence deux résultats fondamentaux.

Tout d'abord, les résultats expérimentaux montrent que le cervelet est fortement impliqué dans des processus linguistiques moteurs et non moteurs très variés. Marien *et al.* (2001a) répartissent ces processus en cinq catégories (voir ci-après) : 1^{o} les mécanismes de mémoire de travail associés aux différents types de tâches linguistiques ; 2^{o} la planification articulatoire ; 3^{o} certaines opérations linguistiques impliquées dans les aspects sémantique et phonologique de la recherche de mots ; 4^{o} le traitement syntaxique ; 5^{o} les aspects dynamiques du traitement du langage.

Ensuite, les recherches indiquent que dans une très grande majorité de ces tâches linguistiques, les activations du cervelet sont latéralisées dans l'hémisphère droit (voir ci-après). Cette

latéralisation cérébelleuse droite est à mettre en correspondance avec la latéralisation prédominante à gauche dans le cerveau pour des tâches linguistiques — en tenant compte du fait que les connexions cérebello-corticales sont croisées. En particulier, il existe des connexions réciproques entre le cortex préfrontal gauche et le bord latéral droit du cervelet (Leiner *et al.*, 1986, 1989; Engelborghs *et al.*, 1998).

Fonctions linguistiques cérébelleuses latéralisées à droite
(entre parenthèses : pathologie associée)

- articulation furtive : Fiez et Raichle (1997)
- planification articulatoire *(apraxie de la parole)* : Marien *et al.* (2001b)
- discrimination du trait de voisement et de la durée des voyelles : Ackermann *et al.* (1997)
- associations verbales *(déficit de recherche lexicale)* : Petersen *et al.* (1988)
- génération de synonymes, de traductions, de rimes *(déficit de recherche lexicale)* : Klein *et al.* (1995)
- complètement de mots *(déficit de recherche lexicale)* : Buckner *et al.* (1995)
- associations sémantiques *(déficit de recherche lexicale)* : Fiez *et al.* (1992)
- génération phonologique *(trouble de l'aisance verbale)* : Leggio *et al.* (1995)
- grammaire expressive *(agrammatisme expressif)* : Silveri *et al.* (1994); Zettin *et al.* (1997)
- connaissance syntaxique *(agrammatisme)* : Marien *et al.* (1996, 2000); Riva (1998); Gasparini *et al.* (1999); Fabbro *et al.* (2000); Riva et Giorgi (2000)
- dynamique du langage *(aphasie dynamique)* : Marien *et al.* (1996, 2000); Riva (1998); Gasparini *et al.* (1999); Fabbro *et al.* (2000); Riva et Giorgi (2000)

Fonctions linguistiques cérébelleuses latéralisées à gauche
(entre parenthèses : pathologie associée)

- contrôle musculaire de la parole *(dysarthrie ataxique)* : Holmes (1917); Lechtenberg et Gilman (1978)
- organisation visuo-spatiale *(dysgraphie afférente)* : Silveri *et al.* (1997, 1999)

4.5.4 Réseau frontal inférieur droit

La partie du cortex correspondant à l'homologue dans l'hémisphère droit de l'aire de Broca, localisée dans le cortex préfrontal inférieur droit, a montré des activations en relation avec différents types de tâches linguistiques et cognitives :

- inhibition volontaire des réponses (processus cognitif impliqué dans l'annulation d'un mouvement programmé volontairement) (Konishi *et al.*, 1998; Garavan *et al.*, 1999; Konishi *et al.*, 1999; Rubia *et al.*, 2000; Menon *et al.*, 2001; Rubia *et al.*, 2001; Bunge *et al.*, 2002; Garavan *et al.*, 2002; Rubia *et al.*, 2003; Aron *et al.*, 2003, 2004) ;
- rappel en mémoire (Moscovitch *et al.*, 1995; Haxby *et al.*, 1996; Courtney *et al.*, 1997; Fletcher *et al.*, 1998; Wagner *et al.*, 1998) ;
- estimation subjective de l'écoulement du temps (Harrington *et al.*, 1998; Rubia *et al.*, 1998; Brunia *et al.*, 2000; Rubia et Smith, 2004; Shergill *et al.*, 2006) ;
- reconnaissance des visages (Haxby *et al.*, 1994) ;

- traitement des informations faciales émotionnelles (George *et al.*, 1993; Hornak *et al.*, 1996; Nakamura *et al.*, 1999) ;
- production silencieuse de parole mentale (Shergill *et al.*, 2000, 2006) ;
- traitement prosodique de la voix (Zatorre *et al.*, 1992; Kawashima *et al.*, 1993; Imaizumi *et al.*, 1997) ;
- compréhension des métaphores (Bottini *et al.*, 1994) ;
- mémoire de travail phonologique (Paulesu *et al.*, 1993) ;

Bibliographie

Ackermann, H., Graber, S., Hertrich, I. et Daum, I. (1997). Categorical speech perception in cerebellar disorders. *Brain Lang.*, 60(2):323–331.

Allison, T., Puce, A. et McCarthy, G. (2000). Social perception from visual cues : role of the STS region. *Trends Cogn. Sci.*, 4(7):267–278.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J. et Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.*, 6(2):115–116.

Aron, A. R., Robbins, T. W. et Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.*, 8(4):170–177.

Belin, P., Fecteau, S. et Bedard, C. (2004). Thinking the voice : neural correlates of voice perception. *Trends Cogn. Sci.*, 8(3):129–135.

Belin, P. et Zatorre, R. J. (2000). "What", "where" and "how" in auditory cortex. *Nat. Neurosci.*, 3(10):965–966.

Belin, P. et Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, 14(16):2105–2109.

Belin, P., Zatorre, R. J. et Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res. Cogn. Brain Res.*, 13(1):17–26.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P. et Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403(6767):309–312.

Bellugi, U., Klima, E. S. et Poizner, H. (1988). Sign language and the brain. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.*, 66:39–56.

Bellugi, U., Poizner, H. et Klima, E. S. (1983). Brain organization for language : clues from sign aphasia. *Hum. Neurobiol.*, 2(3):155–170.

Bernstein, L. E., Auer, Jr., E. T., Moore, J. K., Ponton, C. W., Don, M. et Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, 13:311–315.

Binder, J. R. et Frost, J. A. (1998). Functional MRI studies of language processes in the brain. *Neurosci. News*, 1:15–23.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N. et Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex*, 10(5):512–528.

Bloedel, J. R. et Bracha, V. (1997). Duality of cerebellar motor and cognitive functions. *Int. Rev. Neurobiol.*, 41:613–634.

Boltshauser, E. et Isler, W. (1977). Joubert syndrome : episodic hyperpnea, abnormal eye movements, retardation and ataxia, associated with dysplasia of the cerebellar vermis. *Neuropädiatrie*, 8(1):57–66.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S. et Frith, C. D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain*, 117:6.

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie. *Bull. Soc. Anat. Paris*, 6:330–357.

Brunia, C. H., de Jong, B. M., van den Berg-Lenssen, M. M. et Paans, A. M. (2000). Visual feedback about time estimation is related to a right hemisphere activation measured by PET. *Exp. Brain Res.*, 130(3):328–337.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A. et Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics : an FMRI study. *J. Cogn. Neurosci.*, 16(1):114–126.

Buchsbaum, B., Hickok, G. et Humphries, C. (2001). Role of the left posterior temporal gyrus in phonological processing for speech perception and production. *Cogn. Sci.*, 25:663–678.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R. et Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *J. Neurosci.*, 15(1 Pt 1):12–29.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J. et Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children : evidence from fMRI. *Neuron*, 33(2):301–311.

Callan, D. E., Callan, A. M., Kroos, C. et Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independant component analysis : A single-sweep EEG case study. *Brain Res. Cogn. Brain Res.*, 10:349–353.

Callan, D. E., Jones, J. A., Munhall, K., Callan, A. M., Kroos, C. et Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, 14(17):2213–2218.

Callan, D. E., Jones, J. A., Munhall, K., Kroos, C., Callan, A. M. et Vatikiotis-Bateson, E. (2004). Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *J. Cogn. Neurosci.*, 16(5):805–816.

Calvert, G. A. (2001). Crossmodal processing in the human brain : insights from functional neuroimaging studies. *Cereb. Cortex*, 11(12):1110–1123.

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D. et David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10:2619–2623.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D. et David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276:593–596.

Calvert, G. A. et Campbell, R. (2003). Reading speech from still and moving faces : the neural substrates of visible speech. *J. Cogn. Neurosci.*, 15(1):57–70.

Calvert, G. A., Campbell, R. et Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.*, 10:649–657.

Calvert, G. A. et Lewis, J. W. (2004). Hemodynamic studies of audio-visual interactions. *In* Calvert, G. A., Spence, C. et Stein, B. E., éditeurs : *The Handbook of Multisensory Processes*, pages 483–502. Cambridge, MA : MIT Press.

Campbell, R. et MacSweeney, M. (2004). Neuroimaging studies of cross-modal plasticity and language processing in deaf people. *In* Calvert, G. A., Spence, C. et Stein, B. E., éditeurs : *The Handbook of Multisensory Processes*, pages 773–782. Cambridge, MA : MIT Press.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J. et David, A. S. (2001). Cortical substrates for the perception of face actions : an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Brain Res. Cogn. Brain Res.*, 12(2):233–243.

Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L. et Chollet, F. (1999). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage*, 9(1):135–144.

Champoux, F., Paiement, P., Mercier, C., Lepore, F., Lassonde, M. et Gagne, J. P. (2007). Auditory processing in a patient with a unilateral lesion of the inferior colliculus. *Eur. J. Neurosci.*, 25(1):291–297.

Champoux, F., Tremblay, C., Mercier, C., Lassonde, M., Lepore, F., Gagne, J. P. et Theoret, H. (2006). A role for the inferior colliculus in multisensory speech integration. *Neuroreport*, 17(15):1607–1610.

Cheney, D. L. et Seyfarth, R. M. (1990). *How monkeys see the world*. Chicago, IL : Univ. Chicago Press.

Clavagnier, S., Falchier, A. et Kennedy, H. (2004). Long-distance feedback projections to area V1 : implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.*, 4(2):117–126.

Colin, C., Radeau, M., Soquet, A. et Deltenre, P. (2004). Generalization of the generation of an MMN by illusory McGurk percepts : Voiceless consonants. *Clin. Neurophysiol.*, 115:1989–2000.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colon, F. et Deltenre, P. (2002). Mismatch negativity evoked by the McGurk-MacDonald effect : A phonetic representation within short-term memory. *Clin. Neurophysiol.*, 113:495–506.

Corina, D. P. (1998). The processing of sign language : Evidence from aphasia. *In* Stemmer, B. et Whitaker, H., éditeurs : *Handbook of neurolinguistics*, pages 313–329. New York : Academic Press.

Corina, D. P. et McBurney, S. L. (2001). The neural representation of language in users of American Sign Language. *J. Commun. Disord.*, 34(6):455–471.

Courtney, S. M., Ungerleider, L. G., Keil, K. et Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, 386(6625):608–611.

Damasio, A. R., Tranel, D. et Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annu. Rev. Neurosci.*, 13:89–109.

Damasio, H. (1991). Neuroanatomical correlates of the aphasias. *In* Sarno, M. T., éditeur : *Acquired Aphasia*, pages 45–71. San Diego CA, Academic Press.

Damasio, H. et Damasio, A. R. (1980). The anatomical basis of conduction aphasia. *Brain*, 103(2):337–350.

De Renzi, E., Perani, D., Carlesimo, G. A., Silveri, M. C. et Fazio, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere–an MRI and PET study and a review of the literature. *Neuropsychologia*, 32(8):893–902.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. et Rizzolatti, G. (1992). Understanding motor events : a neurophysiological study. *Exp. Brain Res.*, 91(1):176–180.

Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A. et Frackowiak, R. S. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115:6.

Démonet, J.-F., Fiez, J. A., Paulesu, E., Petersen, S. E. et Zatorre, R. J. (1996). PET studies of phonological processing : A critical reply to Poeppel. *Brain Lang.*, 55(3):352–379.

Démonet, J.-F., Price, C., Wise, R. et Frackowiak, R. S. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks : a positron-emission tomography study in normal human subjects. *Neurosci. Lett.*, 182(1):25–28.

Démonet, J.-F., Thierry, G. et Cardebat, D. (2005). Renewal of the neurophysiology of language : functional neuroimaging. *Physiol. Rev.*, 85(1):49–95.

Eden, G. F., VanMeter, J. W., Rumsey, J. M. et Zeffiro, T. A. (1996). The visual deficit theory of developmental dyslexia. *Neuroimage*, 4(3 Pt 3):108–117.

Engelborghs, S., Marien, P., Martin, J. J. et De Deyn, P. P. (1998). Functional anatomy, vascularisation and pathology of the human thalamus. *Acta Neurol. Belg.*, 98(3):252–265.

Fabbro, F., Moretti, R. et Bava, A. (2000). Language impairments in patients with cerebellar lesions. *J. Neurolinguist.*, 13:173–188.

Fadiga, L., Craighero, L., Buccino, G. et Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles : a TMS study. *Eur. J. Neurosci.*, 15(2):399–402.

Fadiga, L., Fogassi, L., Pavesi, G. et Rizzolatti, G. (1995). Motor facilitation during action observation : a magnetic stimulation study. *J. Neurophysiol.*, 73(6):2608–2611.

Falchier, A., Clavagnier, S., Barone, P. et Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.*, 22:5749–5759.

Fecteau, S., Armony, J. L., Joanette, Y. et Belin, P. (2004a). Is voice processing species-specific in human auditory cortex ? an fMRI study. *Neuroimage*, 23(3):840–848.

Fecteau, S., Armony, J. L., Joanette, Y. et Belin, P. (2004b). Priming of non-speech vocalizations in male adults : the influence of the speaker's gender. *Brain Cogn.*, 55(2):300–302.

Fehrenbach, R. A., Wallesch, C. W. et Claus, D. (1984). Neuropsychologic findings in Friedreich's ataxia. *Arch. Neurol.*, 41(3):306–308.

Fiez, J. A., Petersen, S. E., Cheney, M. K. et Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. A single case study. *Brain*, 115:1.

Fiez, J. A. et Raichle, M. E. (1997). Linguistic processing. *In* Schmahmann, J., éditeur : *International review of neurobiology : The cerebellum and cognition.*, volume 41, pages 233–254. San Diego, CA : Academic Press.

Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. et Dolan, R. J. (1998). The functional roles of prefrontal cortex in episodic memory. II. retrieval. *Brain*, 121:7.

Foxe, J. J. et Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16:419–423.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W. et Murray, M. M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex : An fMRI study. *J. Neurophysiol.*, 88:540–543.

Frost, J. A., Binder, J. R., Springer, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M. et Cox, R. W. (1999). Language processing is strongly left lateralized in both sexes. Evidence from functional MRI. *Brain*, 122(Pt 2):199–208.

Gallese, V., Fadiga, L., Fogassi, L. et Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119:2.

Garavan, H., Ross, T. J., Murphy, K., Roche, R. A. et Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior : inhibition, error detection, and correction. *Neuroimage*, 17(4):1820–1829.

Garavan, H., Ross, T. J. et Stein, E. A. (1999). Right hemispheric dominance of inhibitory control : an event-related functional MRI study. *Proc. Natl. Acad. Sci. U.S.A.*, 96(14):8301–8306.

Gasparini, M., Di Piero, V., Ciccarelli, O., Cacioppo, M. M., Pantano, P. et Lenzi, G. L. (1999). Linguistic impairment after right cerebellar stroke : a case report. *Eur. J. Neurol.*, 6(3):353–356.

George, M. S., Ketter, T. A., Gill, D. S., Haxby, J. V., Ungerleider, L. G., Herscovitch, P. et Post, R. M. (1993). Brain regions involved in recognizing facial emotion or identity : an oxygen-15 PET study. *J. Neuropsychiatry Clin. Neurosci.*, 5(4):384–394.

Ghazanfar, A. A., Smith-Rohrberg, D. et Hauser, M. D. (2001). The role of temporal cues in rhesus monkey vocal recognition : orienting asymmetries to reversed calls. *Brain Behav. Evol.*, 58(3):163–172.

Giard, M.-H. et Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans : A behavioral and electrophysiological study. *J. Cogn. Neurosci.*, 11:473–490.

Goodglass, H. (1992). Diagnosis of conduction aphasia. *In* Kohn, S. E., éditeur : *Conduction Aphasia*, pages 39–40. Hillsdale NJ, Lawrence Erlbaum Associates.

Goodglass, H. (1993). *Understanding Aphasia*. San Diego CA, Academic Press.

Grafton, S. T., Arbib, M. A., Fadiga, L. et Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Exp. Brain Res.*, 112(1):103–111.

Grasby, P. M., Frith, C. D., Friston, K. J., Simpson, J., Fletcher, P. C., Frackowiak, R. S. et Dolan, R. J. (1994). A graded task approach to the functional mapping of brain areas implicated in auditory-verbal memory. *Brain*, 117:6.

Habib, M., Joanette, Y. et Roch Lecours, A. (2000). Le cerveau humain et les origines du langage. *Med. Sci.*, 16(2):171–180.

Hamilton, N. G., Frick, R. B., Takahashi, T. et Hopping, M. W. (1983). Psychiatric symptoms and cerebellar pathology. *Am. J. Psychiatry*, 140(10):1322–1326.

Harrington, D. L., Haaland, K. Y. et Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *J. Neurosci.*, 18(3):1085–1095.

Hauser, M. D. et Andersson, K. (1994). Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys : field experiments. *Proc. Natl. Acad. Sci. U.S.A.*, 91(9):3946–3948.

Haxby, J. V., Hoffman, E. A. et Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.*, 4(6):223–233.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P. et Grady, C. L. (1994). The functional organization of human extrastriate cortex : a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.*, 14(11 Pt 1):6336–6353.

Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I. et Grady, C. L. (1996). Face encoding and recognition in the human brain. *Proc. Natl. Acad. Sci. U.S.A.*, 93(2):922–927.

Heffner, H. E. et Heffner, R. S. (1984). Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science*, 226(4670):75–76.

Hickok, G. (2001). Functional anatomy of speech perception and speech production : psycholinguistic implications. *J. Psycholinguist Res.*, 30(3):225–235.

Hickok, G., Bellugi, U. et Klima, E. S. (1996). The neurobiology of sign language and its implications for the neural basis of language. *Nature*, 381(6584):699–702.

Hickok, G., Bellugi, U. et Klima, E. S. (1998). WhatŠs right about the neural organization of sign language ? A perspective on recent neuroimaging results. *Trends Cogn. Sci.*, 2:465–468.

Hickok, G., Buchsbaum, B., Humphries, C. et Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI : speech, music, and working memory in area spt. *J. Cogn. Neurosci.*, 15(5):673–682.

Hickok, G. et Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.*, 4(4):131–138.

Hickok, G. et Poeppel, D. (2004). Dorsal and ventral streams : a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99.

Hirata, M., Kato, A., Taniguchi, M., Saitoh, Y., Ninomiya, H., Ihara, A., Kishima, H., Oshino, S., Baba, T., Yorifuji, S. et Yoshimine, T. (2004). Determination of language dominance with synthetic aperture magnetometry : comparison with the Wada test. *Neuroimage*, 23(1):46–53.

Holmes, G. (1917). The symptoms of acute cerebellar injuries due to gunshot injuries. *Brain*, 40:461–535.

Hornak, J., Rolls, E. T. et Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*, 34(4):247–261.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. et Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449):2526–2528.

Imaizumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H., Itoh, K., Kato, T., Nakamura, A., Hatano, K., Kojima, S. et Nakamura, K. (1997). Vocal identification of speaker and emotion activates different brain regions. *Neuroreport*, 8(12):2809–2812.

Indefrey, P. et Levelt, W. J. (2000). The neural correlates of language production. *In* Gazzaniga, M. S., éditeur : *The New Cognitive Neurosciences*, pages 845–865. Cambridge, MA : MIT Press.

Jaeger, J. J., Lockwood, A. H., Van Valin, Jr, R. D., Kemmerer, D. L., Murphy, B. W. et Wack, D. S. (1998). Sex differences in brain regions activated by grammatical and reading tasks. *Neuroreport*, 9(12):2803–2807.

Jones, J. A. et Callan, D. E. (2003). Brain activity during audiovisual speech perception : an fMRI study of the McGurk effect. *Neuroreport*, 14(8):1129–1133.

Kang, E., Lee, D. S., Kang, H., Hwang, C. H., Oh, S. H., Kim, C. S., Chung, J. K. et Lee, M. C. (2006). The neural correlates of cross-modal interaction in speech perception during a semantic decision task on sentences : a PET study. *Neuroimage*, 32(1):423–431.

Kansaku, K. et Kitazawa, S. (2001). Imaging studies on sex differences in the lateralization of language. *Neurosci. Res.*, 41(4):333–337.

Kansaku, K., Yamaura, A. et Kitazawa, S. (2000). Sex differences in lateralization revealed in the posterior language areas. *Cereb. Cortex*, 10(9):866–872.

Kanwisher, N., McDermott, J. et Chun, M. M. (1997). The fusiform face area : a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, 17(11):4302–4311.

Kanwisher, N. et Yovel, G. (2006). The fusiform face area : a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 361(1476):2109–2128.

Kawashima, R., Itoh, M., Hatazawa, J., Miyazawa, H., Yamada, K., Matsuzawa, T. et Fukuda, H. (1993). Changes of regional cerebral blood flow during listening to an unfamiliar spoken language. *Neurosci. Lett.*, 161(1):69–72.

Kimura, D. (1973). Manual activity during speaking. II. Left-handers. *Neuropsychologia*, 11(1):51–55.

Kirveskari, E., Salmelin, R. et Hari, R. (2006). Neuromagnetic responses to vowels vs. tones reveal hemispheric lateralization. *Clin. Neurophysiol.*, 117(3):643–648.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E. et Evans, A. C. (1995). The neural substrates underlying word generation : a bilingual functional-imaging study. *Proc. Natl. Acad. Sci. U.S.A.*, 92(7):2899–2903.

Konishi, S., Nakajima, K., Uchida, I., Kikyo, H., Kameyama, M. et Miyashita, Y. (1999). Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. *Brain*, 122:5.

Konishi, S., Nakajima, K., Uchida, I., Sekihara, K. et Miyashita, Y. (1998). No-go dominant brain activity in human inferior prefrontal cortex revealed by functional magnetic resonance imaging. *Eur. J. Neurosci.*, 10(3):1209–1213.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres : a computational approach. *Psychol. Rev.*, 94(2):148–175.

Lechtenberg, R. et Gilman, S. (1978). Speech disorders in cerebellar disease. *Ann. Neurol.*, 3(4):285–290.

Leggio, M. G., Solida, A., Silveri, M. C., Gainotti, G. et Molinari, M. (1995). Verbal fluency impairments in patients with cerebellar lesions. *Soc. Neurosci. Abstr.*, 21:917.

Leiner, H. C., Leiner, A. L. et Dow, R. S. (1986). Does the cerebellum contribute to mental skills? *Behav. Neurosci.*, 100(4):443–454.

Leiner, H. C., Leiner, A. L. et Dow, R. S. (1989). Reappraising the cerebellum : what does the hindbrain contribute to the forebrain ? *Behav. Neurosci.*, 103(5):998–1008.

Levelt, W. J., Praamstra, P., Meyer, A. S., Helenius, P. et Salmelin, R. (1998). An MEG study of picture naming. *J. Cogn. Neurosci.*, 10(5):553–567.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. et Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.*, 74(6):431–461.

Liberman, A. M. et Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1):1–36.

Liberman, A. M. et Whalen, D. H. (2000). On the relation of speech to language. *Trends Cogn. Sci.*, 4(5):187–196.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., Williams, S. C., Woll, B. et Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise : An event-related fMRI study. *Neuroreport*, 11:1729–1733.

MacSweeney, M., Calvert, G. A., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C., Woll, B. et Brammer, M. J. (2002). Speechreading circuits in people born deaf. *Neuropsychologia*, 40(7):801–807.

MacSweeney, M., Campbell, R., Calvert, G. A., McGuire, P. K., David, A. S., Suckling, J., Andrew, C., Woll, B. et Brammer, M. J. (2001). Dispersed activation in the left temporal cortex for speech-reading in congenitally deaf people. *Proc. Biol. Sci.*, 268(1466):451–457.

Marien, P., Engelborghs, S., Fabbro, F. et De Deyn, P. P. (2001a). The lateralized linguistic cerebellum : a review and a new hypothesis. *Brain Lang.*, 79(3):580–600.

Marien, P., Engelborghs, S., Pickut, B. A. et De Deyn, P. P. (2000). Aphasia following cerebellar damage : fact or fallacy ? *J. Neurolinguist.*, 13(2):145–171.

Marien, P., Pickut, B. A., Engelborghs, S., Martin, J. J. et De Deyn, P. P. (2001b). Phonological agraphia following a focal anterior insulo-opercular infarction. *Neuropsychologia*, 39(8):845–855.

Marien, P., Saerens, J., Nanhoe, R., Moens, E., Nagels, G., Pickut, B. A., Dierckx, R. A. et De Deyn, P. P. (1996). Cerebellar induced aphasia : case report of cerebellar induced prefrontal aphasic language phenomena supported by SPECT findings. *J. Neurol. Sci.*, 144(1-2):34–43.

McCarthy, G., Puce, A., Gore, J. C. et Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *J. Cogn. Neurosci.*, 9:605–610.

McGuire, P. K., Robertson, D., Thacker, A., David, A. S., Kitson, N., Frackowiak, R. S. et Frith, C. D. (1997). Neural correlates of thinking in sign language. *Neuroreport*, 8(3):695–698.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H. et Reiss, A. L. (2001). Error-related brain activation during a go/nogo response inhibition task. *Hum. Brain Mapp.*, 12(3):131–143.

Molholm, S. et Foxe, J. J. (2005). Look "hear", primary auditory cortex is active during lip-reading. *Neuroreport*, 16:123–124.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E. et Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans : A high-density electrical mapping study. *Brain Res. Cogn. Brain Res.*, 14:115–128.

Moscovitch, C., Kapur, S., Kohler, S. et Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity : a positron emission tomography study in humans. *Proc. Natl. Acad. Sci. U.S.A.*, 92(9):3721–3725.

Möttönen, R., Jarvelainen, J., Sams, M. et Hari, R. (2005). Viewing speech modulates activity in the left SI mouth cortex. *Neuroimage*, 24(3):731–737.

Möttönen, R., Krause, C. M., Tippana, K. et Sams, M. (2002). Processing of visual speech in the human auditory cortex. *Brain Res. Cogn. Brain Res.*, 13:417–425.

Möttönen, R., Schürmann, M. et Sams, M. (2004). Time course of multisensory interactions during audiovisual speech perception in humans : a magnetoencephalographic study. *Neurosci. Lett.*, 363(2):112–115.

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C. E. et Foxe, J. J. (2005). Grabbing your ear : Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex*, 22:5055–5073.

Musacchia, G., Sams, M., Nicol, T. et Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Exp. Brain Res.*, 168:1–10.

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H. et Kojima, S. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *J. Neurophysiol.*, 82(3):1610–1614.

Neuner, F. et Schweinberger, S. R. (2000). Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain Cogn.*, 44(3):342–366.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P. et Turner, R. (1998). Cerebral organization for language in deaf and hearing subjects : biological constraints and effects of experience. *Proc. Natl. Acad. Sci. U.S.A.*, 95(3):922–929.

Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. et Kubo, T. (1999). Sign language 'heard' in the auditory cortex. *Nature*, 397(6715):116.

Nishitani, N. et Hari, R. (2002). Viewing lip forms : cortical dynamics. *Neuron*, 36(6):1211–1220.

Okada, K. et Hickok, G. (2006). Left posterior auditory-related cortices participate both in speech perception and speech production : neural overlap revealed by fMRI. *Brain Lang.*, 98(1):112–117.

Okada, K., Smith, K. R., Humphries, C. et Hickok, G. (2003). Word length modulates neural activity in auditory cortex during covert object naming. *Neuroreport*, 14(18):2323–2326.

Olson, I. R., Gatenby, J. C. et Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Brain Res. Cogn. Brain Res.*, 14(1):129–138.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B. et Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production : a contribution to the definition of language epicenters with PET. *Neuroimage*, 11(4):347–357.

Parrot, M., Doyon, B., Démonet, J.-F. et Cardebat, D. (1999). Hemispheric preponderance in categorical and coordinate visual processes. *Neuropsychologia*, 37(11):1215–1225.

Paulesu, E., Frith, C. D. et Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362(6418):342–345.

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S. et Frith, C. D. (1996). Is developmental dyslexia a disconnection syndrome? evidence from PET scanning. *Brain*, 119:1.

Paulesu, E. et Mehler, J. (1998). Right on in sign language. *Nature*, 392(6673):233–234.

Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., Sensolo, S. et Fazio, F. (2003). A functional-anatomical model for lipreading. *J. Neurophysiol.*, 90(3):2005–2013.

Pekkola, J., Laasonen, M., Ojanen, V., Autti, T., Jääskeläinen, I. P., Kujala, T. et Sams, M. (2006). Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers : an fMRI study at 3 T. *Neuroimage*, 29(3):797–807.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A. et Sams, M. (2005). Primary auditory cortex activation by visual speech : An fMRI study at 3T. *Neuroreport*, 16:125–128.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. et Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331(6157):585–589.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D. et Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages : implications for the neural basis of human language. *Proc. Natl. Acad. Sci. U.S.A.*, 97(25):13961–13966.

Phillips, M. D., Lowe, M. J., Lurito, J. T., Dzemidzic, M. et Mathews, V. P. (2001). Temporal lobe activation demonstrates sex-based differences during passive listening. *Radiology*, 220(1):202–207.

Poizner, H., Bellugi, U. et Klima, E. S. (1990). Biological foundations of language : clues from sign language. *Annu. Rev. Neurosci.*, 13:283–307.

Poizner, H., Klima, E. et Bellugi, U. (1987). *What the hands reveal about the brain.* Cambridge, MA : MIT Press.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H. et Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, 10(1):15–35.

Poremba, A., Malloy, M., Saunders, R. C., Carson, R. E., Herscovitch, P. et Mishkin, M. (2004). Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature*, 427(6973):448–451.

Prell, C. G., Hauser, M. D. et Moody, D. B. (2002). Discrete or graded variation within rhesus monkey screams? Psychophysical experiment on classification. *Anim. Behav.*, 63:47.

Price, C. J., Mummery, C. J., Moore, C. J., Frackowiak, R. S. et Friston, K. J. (1999). Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients. *J. Cogn. Neurosci.*, 11(4):371–382.

Puce, A., Allison, T., Bentin, S., Gore, J. C. et McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.*, 18(6):2188–2199.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Shankweiler, D. P., Katz, L., Fletcher, J. M. et Gore, J. C. (1996). Cerebral organization of component processes in reading. *Brain*, 119:4.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O. et Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.*, 103(20):7865–7870.

Rasmussen, T. et Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Ann. N.Y. Acad. Sci.*, 299:355–369.

Reale, R. A., Calvert, G. A., Thesen, T., Jenison, R. L., Kawasaki, H., Oya, H., Howard, M. A. et Brugge, J. F. (2007). Auditory-visual processing represented in the human superior temporal gyrus. *Neuroscience*, 145(1):162–184.

Riva, D. (1998). The cerebellar contribution to language and sequential functions : evidence from a child with cerebellitis. *Cortex*, 34(2):279–287.

Riva, D. et Giorgi, C. (2000). The cerebellum contributes to higher functions during development : evidence from a series of children surgically treated for posterior fossa tumours. *Brain*, 123:5.

Rizzolatti, G. et Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.*, 21(5):188–194.

Rizzolatti, G., Fadiga, L., Gallese, V. et Fogassi, L. (1996a). Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.*, 3(2):131–141.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D. et Fazio, F. (1996b). Localization of grasp representations in humans by PET : 1. Observation versus execution. *Exp. Brain Res.*, 111(2):246–252.

Rönnberg, J., Söderfeldt, B. et Risberg, J. (2000). The cognitive neuroscience of signed language. *Acta Psychol.*, 105(2-3):237–254.

Rockland, K. S. et Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.*, 50(1-2):19–26.

Rossell, S. L., Bullmore, E. T., Williams, S. C. et David, A. S. (2002). Sex differences in functional brain activation during a lexical visual field task. *Brain Lang.*, 80(1):97–105.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S., Simmons, A., Andrew, C. et Bullmore, E. (1998). Prefrontal involvement in "temporal bridging" and timing movement. *Neuropsychologia*, 36(12):1283–1293.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S. C., Simmons, A., Andrew, C. et Bullmore, E. T. (2000). Functional frontalisation with age : mapping neurodevelopmental trajectories with fMRI. *Neurosci. Biobehav. Rev.*, 24(1):13–19.

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., Simmons, A., Williams, S. C., Giampietro, V., Andrew, C. M. et Taylor, E. (2001). Mapping motor inhibition : conjunctive brain activations across different versions of go/no-go and stop tasks. *Neuroimage*, 13(2):250–261.

Rubia, K. et Smith, A. (2004). The neural correlates of cognitive time management : a review. *Acta Neurobiol. Exp.*, 64(3):329–340.

Rubia, K., Smith, A. B., Brammer, M. J. et Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage*, 20(1):351–358.

Ruff, S., Cardebat, D., Marie, N. et Démonet, J.-F. (2002). Enhanced response of the left frontal cortex to slowed down speech in dyslexia : an fMRI study. *Neuroreport*, 13(10):1285–1289.

Rutten, G. J., Ramsey, N. F., van Rijen, P. C., Alpherts, W. C. et van Veelen, C. W. (2002). fMRI-determined language lateralization in patients with unilateral or mixed language dominance according to the Wada test. *Neuroimage*, 17(1):447–460.

Ruytjens, L., Albers, F., van Dijk, P., Wit, H. et Willemsen, A. (2006). Neural responses to silent lipreading in normal hearing male and female subjects. *Eur. J. Neurosci.*, 24(6):1835–1844.

Sabbah, P., Chassoux, F., Leveque, C., Landre, E., Baudoin-Chial, S., Devaux, B., Mann, M., Godon-Hardy, S., Nioche, C., Ait-Ameur, A., Sarrazin, J. L., Chodkiewicz, J. P. et Cordoliani, Y. S. (2003). Functional MR imaging in assessment of language dominance in epileptic patients. *Neuroimage*, 18(2):460–467.

Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W. et Foxe, J. J. (2007). Seeing voices : High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia*, 45(3):587–597.

Salmelin, R., Hari, R., Lounasmaa, O. V. et Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature*, 368(6470):463–465.

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T. et Simola, J. (1991). Seeing speech : visual information from lip movements modifies activity in the human auditory cortex. *Neurosci. Lett.*, 127:141–145.

Sams, M., Manninen, P., Surakka, V., Helin, P. et Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences : Effects of word meaning and sentence context. *Speech Comm.*, 26:75–87.

Schirmer, A., Kotz, S. A. et Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Res. Cogn. Brain Res.*, 14(2):228–233.

Schirmer, A., Striano, T. et Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, 16(6):635–639.

Schroeder, C. E. et Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.*, 14(1):187–198.

Schroeder, C. E. et Foxe, J. J. (2005). Multisensory contributions to low-level "unisensory" processing. *Curr. Opin. Neurobiol.*, 15:454–458.

Schwartz, J.-L. (2004). La parole multisensorielle : plaidoyer, problèmes, perspective. *In Journées d'étude sur la parole*. Fès, Maroc.

Scott, S. K., Blank, C. C., Rosen, S. et Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123:12.

Söderfeldt, B., Ingvar, M., Rönnberg, J., Eriksson, L., Serrander, M. et Stone-Elander, S. (1997). Signed and spoken language perception studied by positron emission tomography. *Neurology*, 49(1):82–87.

Söderfeldt, B., Rönnberg, J. et Risberg, J. (1994). Regional cerebral blood flow in sign language users. *Brain Lang.*, 46(1):59–68.

Sekiyama, K., Kanno, I., Miura, S. et Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci. Res.*, 47(3):277–287.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweiler, D. P., Katz, L. et Gore, J. C. (1995). Sex differences in the functional organization of the brain for language. *Nature*, 373(6515):607–609.

Shergill, S. S., Bullmore, E., Simmons, A., Murray, R. et McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *Am. J. Psychiatry*, 157(10): 1691–1693.

Shergill, S. S., Tracy, D. K., Seal, M., Rubia, K. et McGuire, P. (2006). Timing of covert articulation : an fMRI study. *Neuropsychologia*, 44(12):2573–2577.

Silveri, M. C., Leggio, M. G. et Molinari, M. (1994). The cerebellum contributes to linguistic production : a case of agrammatic speech following a right cerebellar lesion. *Neurology*, 44(11):2047–2050.

Silveri, M. C., Misciagna, S., Leggio, M. G. et Molinari, M. (1997). Spatial dysgraphia and cerebellar lesion : a case report. *Neurology*, 48(6):1529–1532.

Silveri, M. C., Misciagna, S., Leggio, M. G. et Molinari, M. (1999). Cerebellar spatial dysgraphia : further evidence. *J. Neurol.*, 246(4):312–313.

Skipper, J. I., Nusbaum, H. C. et Small, S. L. (2005). Listening to talking faces : motor cortical activation during speech perception. *Neuroimage*, 25(1):76–89.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C. et Small, S. L. (2007). Hearing lips and seeing voices : how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex.*

Strelnikov, K. (2007). Activation-verification in continuous speech processing. Interaction of cognitive strategies as a possible theoretical approach. *J. Neurolinguist.*

Sundara, M., Namasivayam, A. K. et Chen, R. (2001). Observation-execution matching system for speech : a magnetic stimulation study. *Neuroreport*, 12(7):1341–1344.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.*, 9(2):182–198.

Tallal, P., Miller, S. et Fitch, R. H. (1993). Neurobiological basis of speech : a case for the preeminence of temporal processing. *Ann. N.Y. Acad. Sci.*, 682:27–47.

Van Lancker, D. R., Cummings, J. L., Kreiman, J. et Dobkin, B. H. (1988). Phonagnosia : a dissociation between familiar and unfamiliar voices. *Cortex*, 24(2):195–209.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B. et Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas : phonology, semantics, and sentence processing. *Neuroimage*, 30(4):1414–1432.

von Kriegstein, K. et Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2):948–955.

von Kriegstein, K., Kleinschmidt, A. et Giraud, A.-L. (2006). Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cereb. Cortex*, 16(9):1314–1322.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P. et Giraud, A.-L. (2005). Interaction of face and voice areas during speaker recognition. *J. Cogn. Neurosci.*, 17(3):367–376.

Wagner, A. D., Poldrack, R. A., Eldridge, L. L., Desmond, J. E., Glover, G. H. et Gabrieli, J. D. (1998). Material-specific lateralization of prefrontal activation during episodic encoding and retrieval. *Neuroreport*, 9(16):3711–3717.

Watkins, K. et Paus, T. (2004). Modulation of motor excitability during speech perception : the role of Broca's area. *J. Cogn. Neurosci.*, 16(6):978–987.

Watkins, K. E., Strafella, A. P. et Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8):989–994.

Wernicke, C. (1874). *Der aphasiche Symptomencomplex*. Breslau : Cohn and Weigert.

Wilshire, C. E. et McCarthy, R. A. (1996). Experimental investigations of an impairment in phonological encoding. *Cogn. Neuropsychol.*, 13(7):1059–1098.

Wilson, S. M. et Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility : evidence for the sensorimotor nature of speech perception. *Neuroimage*, 33(1):316–325.

Wilson, S. M., Saygin, A. P., Sereno, M. I. et Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.*, 7(7):701–702.

Wise, R. J., Greene, J., Buchel, C. et Scott, S. K. (1999). Brain regions involved in articulation. *Lancet*, 353(9158):1057–1061.

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K. et Warburton, E. A. (2001). Separate neural subsystems within "Wernicke's area". *Brain*, 124(Pt 1):83–95.

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J. et McCarthy, G. (2003). Polysensory interactions along lateral regions evoked by audiovisual speech. *Cereb. Cortex*, 13:1034.

Zatorre, R. J., Belin, P. et Penhune, V. B. (2002). Structure and function of auditory cortex : music and speech. *Trends Cogn. Sci.*, 6(1):37–46.

Zatorre, R. J., Evans, A. C., Meyer, E. et Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058):846–849.

Zatorre, R. J., Meyer, E., Gjedde, A. et Evans, A. C. (1996). PET studies of phonetic processing of speech : review, replication, and reanalysis. *Cereb. Cortex*, 6(1):21–30.

Zettin, M., Cappa, S. F., D'Amico, A., Rago, R., Perino, C., Perani, D. et Fazio, F. (1997). Agrammatic speech production after a right cerebellar haemorrhage. *Neurocase*, 3:375–380.

Chapitre 5

Intégration multisensorielle

Sommaire

5.1 Multisensorialité : principes et propriétés 110
5.2 Multisensorialité et congruence spatio-temporelle 110
5.2.1 Congruence spatiale 111
5.2.2 Congruence temporelle 112
5.3 Autres formes d'interaction multisensorielle 113
5.3.1 Perception de la parole 113
5.3.2 Perception du mouvement 113
5.3.3 Perception des goûts et des odeurs 113
5.4 Substrats neuronaux de la perception multisensorielle 114
5.4.1 Intégration dans les aires corticales associatives 114
5.4.2 Intégration dans les aires corticales de bas-niveau 116
5.4.3 Intégration dans les régions sous-corticales 117
Bibliographie 119

Introduction

Les perceptions et les comportements animaux sont déterminés par l'intégration des informations en provenance de différentes sources sensorielles. La perception de la parole, activité fondamentalement audiovisuelle, est un exemple canonique d'intégration multisensorielle faisant intervenir les modalités auditive et visuelle (voir *« Perception audiovisuelle de la parole »*, ch. 3, p. 59 et *« Intégration neuronale de la parole et du langage »*, ch. 4, p. 79). Les conceptions scientifiques de la sensorialité se sont adaptées progressivement pour prendre pleinement en compte cette réalité : alors que les recherches neurobiologiques pionnières envisagaient les sensorialités de façon indépendante et compartimentée, des premières études sur la multisensorialité ont vu le jour dans la première moitié du XX^{e} siècle et les connaissances dans le domaine n'ont depuis plus cessé de progresser.

En parallèle, la vision scientifique de la multisensorialité a subi une transition en passant d'une organisation très hiérarchisée, plaçant l'intégration multisensorielle au bout d'une chaîne de traitements unisensoriels spécifiques à chaque modalité, à une organisation beaucoup plus souple et complexe, introduisant l'intégration multisensorielle dès les

premières étapes de traitement des informations perceptives. Ainsi les découvertes les plus récentes démontrent que la notion d'aire corticale primaire (c'est-à-dire strictement circonscrite à une modalité sensorielle particulière) doit à l'avenir être utilisée avec beaucoup de parcimonie, le traitement multisensoriel des perceptions se produisant déjà en partie dans les structures sous-corticales.

5.1 Multisensorialité : principes et propriétés

Stein et Meredith (1993) ont explicité plusieurs principes régissant de façon générale les mécanismes d'intégration multisensorielle. Le principe le plus important est celui de la *coïncidence spatio-temporelle* : les signaux en provenance des différentes modalités sensorielles sont intégrés ensemble de façon optimale lorsqu'ils coïncident à la fois temporellement (ils sont perçus en même temps) et spatialement (ils sont perçus comme provenant de la même direction). On verra dans la suite que ce principe de coïncidence doit être assoupli en principe de *congruence* dans certaines situations particulières.

Un second principe essentiel est celui d'*efficacité inverse* ("inverted effectiveness") : l'efficacité relative de l'intégration multisensorielle est d'autant plus marquée que les signaux unimodaux sont difficiles à détecter. Il ne faut toutefois pas oublier que cette efficacité inverse est limitée aux cas où la combinaison des différentes modalités sensorielles fournit encore suffisamment d'informations pour constituer un percept ; en effet, en-deçà d'un certain seuil perceptif, des signaux unimodaux difficiles à détecter ne seront pas mieux perçus après intégration multisensorielle. Ainsi, entre le cas le plus favorable, où les stimuli unimodaux sont si faciles à détecter que la multisensorialité ne permet aucune amélioration perceptive, et le cas le plus défavorable, où les stimuli unimodaux sont si difficiles à détecter que la multisensorialité n'apporte pas plus d'informations perceptives, on peut s'attendre à la présence systématique d'un seuil perceptif intermédiaire pour lequel le gain multisensoriel est optimal.

5.2 Multisensorialité et congruence spatio-temporelle

Lorsque les informations issues de plusieurs modalités sensorielles sont synchrones (coïncidence temporelle) et proviennent de la même direction de l'espace (coïncidence spatiale), on observe une facilitation comportementale pour la détection et l'orientation vers la source perceptive (Stein *et al.*, 1988, 1989; Wilkinson *et al.*, 1996; Jiang *et al.*, 2002), ainsi qu'une amélioration du temps de réaction en comparaison avec une stimulation unimodale (Bernstein *et al.*, 1969; Zahn *et al.*, 1978; Gielen *et al.*, 1983; Engelken et Stevens, 1989; Perrott *et al.*, 1990; Hughes *et al.*, 1994; Frens *et al.*, 1995; Corneil et Munoz, 1996; Goldring *et al.*, 1996; Frens et Van Opstal, 1998; Harrington et Peck, 1998; Giard et Peronnet, 1999).

Avant de nous intéresser aux effets de la congruence spatiale et temporelle lors de l'intégration multisensorielle, explicitons la différence entre congruence et coïncidence pour éviter toute confusion : alors que la coïncidence, spatiale ou temporelle, est une identité parfaite entre les latences ou les localisations caractérisant les différentes modalités sensorielles, la congruence admet une acception plus large et peut se définir comme une relation constante liant les latences ou les localisations associées aux différentes modalités sensorielles. Ainsi une coïncidence n'est qu'un cas particulier de congruence, c'est-à-dire le cas où la relation qui lie les différentes latences ou localisations est une relation d'identité parfaite. Dans le cas où il existe, par exemple, un décalage constant entre les localisations des sources auditive et visuelle, on peut parler de congruence audiovisuelle ; dans ce cas il n'y a toutefois pas coïncidence. Enfin, dans le cas où il

n'existe aucune relation entre les différentes modalités, les stimuli associés étant complètement indépendants les uns des autres, il n'y a ni coïncidence ni congruence.

5.2.1 Congruence spatiale

Précision spatiale de l'intégration multisensorielle

Les études comportementales examinant de façon détaillée les effets de la congruence spatiale sur l'intégration multisensorielle ont principalement été effectuées en combinant les modalités auditive et visuelle. L'une des premières études appliquant la théorie de détection du signal (SDT : signal detection theory) à la congruence spatiale audiovisuelle a démontré une facilitation perceptive visuelle par un stimulus auditif émis dans le même hémichamp (McDonald *et al.*, 2000). Une étude ultérieure basée sur les mêmes principes mais employant un découpage spatial très détaillé a démontré que la facilitation perceptive audiovisuelle ne se produisait que pour des séparations angulaires inférieures à 8° dans le plan horizontal (Frassinetti *et al.*, 2002). Dans ces deux cas, l'utilisation de la théorie du signal a permis de confirmer que cette facilitation perceptive était bien due à une diminution effective du seuil de détectabilité.

Effets de la disparité spatiale

Dans le cas où il n'y pas coïncidence spatiale entre les stimuli en provenance des différentes modalités, la disparité spatiale ainsi provoquée induit une inhibition réciproque entre les différents stimuli (Stein *et al.*, 1988, 1989; Meredith et Stein, 1996; Kadunce *et al.*, 1997; Jiang *et al.*, 2002).

L'un des effets les plus remarquables de la disparité spatiale sur l'intégration multisensorielle est celui de la ventriloquie (Howard et Templeton, 1966; Bertelson, 1998), une illusion audiovisuelle couramment employée dans le monde du spectacle. Le principe est le suivant : un marionnettiste s'entraîne à parler sans mouvement apparent des lèvres, tout en synchronisant les mouvements de la bouche de sa marionnette sur les paroles émises ; ainsi le public a l'impression que le son sort de la bouche de la marionnette. Les effets de cette disparité spatiale audiovisuelle ont été testés et étendus sous différentes formes, aussi bien dans des situations reproduisant réalistement l'effet de ventriloquie (Witkin *et al.*, 1952) que dans des situations expérimentales simplifiées utilisant des flashes lumineux et des sons brefs (Choe *et al.*, 1975; Bertelson et Radeau, 1981; Radeau et Bertelson, 1987; Bertelson et Aschersleben, 1998; Bertelson *et al.*, 2000; Vroomen *et al.*, 2001; Alais et Burr, 2004). Dans tous les cas on observe une erreur de localisation de la source sonore, en particulier lorsque le décalage spatial entre les sources auditive et visuelle est inférieur ou proche du seuil de précision en localisation auditive (Slutsky et Recanzone, 2001).

Dans ce cas, c'est la vision qui provoque une illusion spatiale auditive, contrairement à l'illusion temporelle créée par Shams *et al.* (2000) (voir « *Congruence temporelle* », § 5.2.2, p. 112), où c'était l'audition qui provoquait une illusion visuelle. Cette différence peut être mise en parallèle avec les performances de discrimination spatiale, qui sont meilleures pour la vision que pour l'audition, ce qui explique que le percept spatial soit fortement influencé par le stimulus visuel dans l'illusion du ventriloque. Des expériences ultérieures ont en outre démontré qu'en cas de disparité spatiale, l'intégration audiovisuelle de la direction de la source est quasi-optimale au sens statistique (Alais et Burr, 2004).

5.2.2 Congruence temporelle

Précision temporelle de l'intégration multisensorielle

Comme dans le cas de la congruence spatiale, les études comportementales examinant en détail les effets de la congruence temporelle sur l'intégration multisensorielle ont pour l'essentiel été réalisées en associant les modalités auditive et visuelle. Ces études ont démontré une facilitation perceptive accrue lorsqu'un stimulus visuel est présenté plusieurs dizaines de millisecondes avant le stimulus auditif correspondant (Hershenson, 1962; Miller, 1986; Giray et Ulrich, 1993; Slutsky et Recanzone, 2001; Corneil *et al.*, 2002). De même, dans les tâches demandant aux sujets une détection explicite du décalage temporel entre audition et vision, cette détection s'avère plus difficile lorsque le stimulus visuel est présenté avant le stimulus auditif associé, tandis qu'elle devient plus aisée dans le cas où le stimulus auditif précède le stimulus visuel (Dixon et Spitz, 1980; Lewkowicz, 1996; Bushara *et al.*, 2001).

Il est intéressant de noter que cette asymétrie temporelle audiovisuelle est plus marquée dans le cas de stimuli de parole, les sujets montrant nettement plus de difficultés pour détecter l'asynchronie lorsque la composante visuelle de la parole est présentée légèrement avant la composante auditive (Dixon et Spitz, 1980). Par ailleurs, cette sensibilité à la synchronie audiovisuelle apparaît très tôt dans le développement humain, dès l'âge de deux mois (Lewkowicz, 1996, 2000).

Cette asymétrie peut être expliquée simplement en considérant la différence entre les vitesses de propagation de la lumière (3×10^8 m.s^{-1}) et du son (3×10^2 m.s^{-1}) dans l'atmosphère. Le temps de propagation de la lumière reste inférieur à 50 µs jusqu'à 10 km de distance et peut donc être négligé dans la plupart des situations écologiques. Au contraire, le temps de propagation du son dépasse 10 ms au-delà de 3 m et atteint 50 ms au-delà de 15 m, de sorte que dans beaucoup de situations écologiques, la composante visuelle présente une avance de plusieurs dizaines de millisecondes sur la composante auditive. La légère asymétrie temporelle observée lors de l'intégration audiovisuelle est donc parfaitement cohérente avec les propriétés physiques qui différencient le décours temporel des stimuli auditifs et visuels ; elle indique une adaptation des circuits multisensoriels aux caractéristiques physiques de l'environnement perceptif.

Effets de la disparité temporelle

Un effet remarquable de la disparité temporelle sur l'intégration multisensorielle peut être observé dans le cas de stimuli audiovisuels incongruents temporellement : dans une tâche visuelle où les sujets doivent comptabiliser une succession rapide de flashes lumineux émis à intervalles constants, l'émission d'une succession de bips auditifs à un rythme différent perturbe l'énumération des flashes visuels (Shams *et al.*, 2000).

Dans ce cas, c'est l'audition qui provoque une illusion temporelle visuelle, contrairement au cas de la ventriloquie (voir *« Congruence spatiale »*, § 5.2.1, p. 111), où c'était la vision qui provoquait une illusion auditive. Cette différence peut être mise en parallèle avec les performances de discrimination temporelle, qui sont meilleures pour l'audition que pour la vision, ce qui explique que le percept temporel soit fortement influencé par le stimulus auditif dans l'illusion de Shams *et al.* (2000). Des expériences ultérieures ont démontré, grâce aux méthodes de l'analyse statistique bayesienne, que le nombre de flashes perçus correspond à celui qui serait perçu par un observateur idéal (au sens de l'optimalité statistique) (Shams *et al.*, 2005).

5.3 Autres formes d'interaction multisensorielle

5.3.1 Perception de la parole

La parole est un stimulus environnemental audiovisuel qui revêt une importance primordiale dans notre vie quotidienne. Dans ce cas particulier on retrouve la notion de congruence audiovisuelle, puisqu'il existe une forte association spatiale, temporelle et phonétique entre, d'une part, les mouvements des lèvres et du visage pour la composante visuelle de la parole, et d'autre part, les variations spectro-temporelles des sons de la voix pour la composante auditive de la parole. On peut donc s'attendre à ce que les composantes auditive et visuelle de la parole interagissent au niveau comportemental et neuronal ; ces thématiques ont été traitées de façon plus détaillée dans *« Perception audiovisuelle de la parole »*, ch. 3, p. 59 et *« Réseau cérébral audio-visuo-moteur de la parole »*, sect. 4.4, p. 84.

Il existe plusieurs façons de perturber l'intégration audiovisuelle de la parole :

- en introduisant un décalage spatial artificiel entre les composantes auditive et visuelle, comme c'est le cas dans les expérience de type ventriloquie (voir plus haut *« Congruence spatiale »*, § 5.2.1, p. 111 ; voir également *« Incongruence spatiale de la parole »*, § 3.4.2, p. 65) ;
- en introduisant un décalage temporel (voir *« Incongruence temporelle de la parole »*, § 3.4.1, p. 64 et *« Congruence temporelle de la parole »*, § 3.3.1, p. 62) ;
- en introduisant une incongruence phonétique entre le phonème vu et le phonème entendu, il s'agit alors de l'effet McGurk (voir *« Conséquences d'une incongruence phonétique : effet McGurk »*, sect. 3.5, p. 65 et *« Congruence phonétique de la parole : complémentarité audiovisuelle »*, § 3.3.2, p. 63).

Dans tous les cas, l'introduction artificielle de disparités spatiales, temporelles ou phonétiques perturbe la détection de la parole (et la modifie considérablement dans le cas de l'effet McGurk), ce qui démontre la forte sensibilité à la congruence audiovisuelle de la parole.

5.3.2 Perception du mouvement

La vision de deux disques se déplaçant dans des directions opposées pour se rapprocher puis s'éloigner peut être interprétée de deux façons différentes : soit on suppose que les disques rebondissent l'un sur l'autre et donc repartent chacun dans la direction opposée à la direction initiale, soit on suppose que les disques se croisent sans interagir et sans changer de direction. En condition normale, une majorité de sujets privilégient l'hypothèse du « croisement ». Toutefois, si un son bref est émis au moment où les disques sont superposés, le percept dominant s'inverse et c'est l'hypothèse du « rebond » qui est majoritairement rapportée (Sekuler *et al.*, 1997). Ce phénomène peut s'expliquer simplement dans la mesure où dans une situation environnementale réelle, deux objets qui se croisent sans se toucher n'émettent pas de bruit, tandis qu'un contact provoque l'émission d'un bruit d'impact. Cette expérience a pu également être répliquée en utilisant un flash lumineux ou une vibration tactile appliquée sur le doigt des sujets (Watanabe et Shimojo, 1998).

5.3.3 Perception des goûts et des odeurs

La coloration artificielle de boissons est susceptible de modifier le goût et l'odeur perçus. Ainsi le goût d'une boisson aromatisée à la cerise est correctement identifié comme de la cerise si cette boisson est colorée en rouge, mais incorrectement identifié comme de l'orange si cette boisson est colorée en orange (Dubose *et al.*, 1980). De façon similaire, l'odeur d'un vin blanc

artificiellement coloré en rouge est décrite avec des adjectifs caractéristiques des vins rouges (Morrot *et al.*, 2001). Ces illusions olfactives et gustatives seraient en partie liées à un effet d'attente perceptive provoquée par la couleur de la boisson.

5.4 Substrats neuronaux de la perception multisensorielle

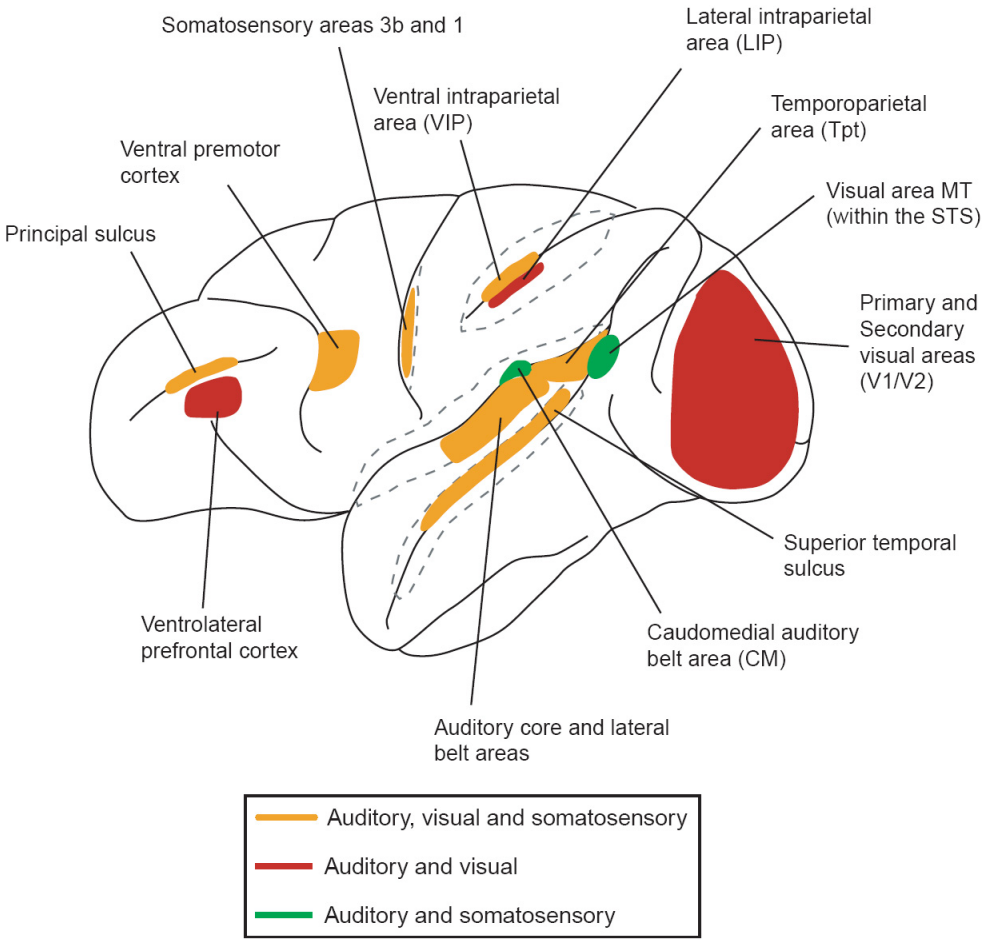


FIG. 5.1 – Anatomie corticale des aires multisensorielles chez le singe. Adapté d'après Ghazanfar et Schroeder (2006).

5.4.1 Intégration dans les aires corticales associatives

Un certain nombre d'études réalisés par électrophysiologie ou par imagerie fonctionnelle ont permis d'identifier plusieurs aires corticales préfrontales, pariétales ou temporales chez l'homme et le singe comme le lieu de « fusion » des différentes modalités sensorielles. Nous dressons dans la suite une revue synoptique de ces différentes aires.

Réseau occipito-temporal

De nombreuses recherches ont démontré l'existence de convergences multisensorielles dans le sulcus temporal supérieur (STS) chez l'homme et dans la région correspondante chez le singe, l'aire polysensorielle temporale supérieure (STP) (Ghazanfar et Schroeder, 2006).

Ainsi, chez le singe, des enregistrements neuronaux directs dans l'aire STP ont montré qu'une proportion importante des cellules nerveuses répondent à la fois aux stimuli auditifs et visuels (Benevento *et al.*, 1977; Bruce *et al.*, 1981; Hikosaka, 1993; Schroeder et Foxe, 2002). Ces études ne montraient pas toutefois pas qu'il y avait réellement intégration des signaux auditifs et visuels. Une étude récente utilisant des stimuli audiovisuels biologiques pertinents, comme des vocalisations de congénères, un expérimentateur humain en train de déchirer du papier, ou en train de marcher, a démontré que la réponse de certains neurones de l'aire STP lors de la vision de ces actions biologiques était modulée par la composante auditive correspondante, cette modulation étant dépendante de la congruence entre les parties auditive et visuelle du stimulus (Barraclough *et al.*, 2005).

Réseau pariétal

La région intrapariétale fait partie d'un réseau cortical important impliqué dans la commande des mouvements dans l'espace en fonction de l'environnement multisensoriel, ce réseau incluant en particulier l'aire intrapariétale latérale (LIP) (Andersen *et al.*, 1997). Longtemps considérée comme une aire unisensorielle visuelle, l'aire LIP comporte en réalité des neurones multisensoriels recevant des informations à la fois visuelles et auditives, ainsi que des informations sur la position des yeux. Chez le singe, lors de tâches en saccades oculaires retardées, les réponses des neurones de l'aire LIP sont modulées par la latence et la position des indices visuels et auditifs (Mazzoni *et al.*, 1996). L'aire intrapariétale ventrale (VIP), adjacente à l'aire intrapariétale latérale, comporte également des neurones multisensoriels répondant à des stimuli visuels, auditifs, somatosensoriels et vestibulaires, les champs récepteurs d'un même neurone dans les différentes modalités se chevauchant la plupart du temps (Duhamel *et al.*, 1998; Bremmer *et al.*, 2002; Avillac *et al.*, 2005; Schlack *et al.*, 2005). La jonction temporo-pariétale (aire Tpt chez le singe), situé au confluent des cortex visuel, auditif et somatosensoriel, joue également un rôle dans la représentation multisensorielle de l'espace péripersonnel. Elle contient par ailleurs des neurones trimodaux dont les champs récepteurs sont situés au niveau de la tête, du cou et des épaules, ce qui indique un rôle possible dans l'orientation de la tête dans l'espace (Leinonen *et al.*, 1980).

Réseau frontal

Le cortex préfrontal est également susceptible de réponses multisensorielles, des études chez le singe ayant révélé l'existence de neurones répondant à la fois aux stimuli auditifs et visuels (Fuster *et al.*, 2000). Le cortex prémoteur contient des neurones trimodaux répondant aux stimuli auditifs, visuels et somatosensoriels, les champs récepteurs correspondants étant principalement localisés au niveau du visage, du torse et des bras (Graziano *et al.*, 1994; Fogassi *et al.*, 1996; Graziano *et al.*, 1999). Chez le singe, ces neurones ont probablement des fonctions liées aux comportements de défense lorsqu'un danger potentiel approche, comme le montrent des études en stimulation sensorielle (Schiff *et al.*, 1962; Ghazanfar *et al.*, 2002; Maier *et al.*, 2004) et en microstimulation (Graziano *et al.*, 2002).

5.4.2 Intégration dans les aires corticales de bas-niveau

De nombreuses études mettent en évidence au niveau des cortex unisensoriels (en particulier les cortex auditif et visuel) une intégration d'informations en provenance d'autres modalités sensorielles, aussi bien chez l'homme que chez l'animal ; par ailleurs cette intégration est même présente dans les aires sensorielles primaires.

Traitement visuel et somatosensoriel dans les aires auditives

Chez l'homme, l'idée d'interactions multisensorielles dès les premières étapes du traitement auditif est soutenue par plusieurs études en potentiels évoqués, démontrant l'existence de courtes latences audiovisuelles (Giard et Peronnet, 1999; Molholm *et al.*, 2002; Besle *et al.*, 2004; van Wassenhove *et al.*, 2005) et audiosomesthésiques (Foxe *et al.*, 2000; Murray *et al.*, 2004). La localisation précise d'interactions multisensorielles dans le planum temporal supérieur a été obtenue dans des études en neuroimagerie bénéficiant d'une bonne résolution spatiale, aussi bien en imagerie par résonance magnétique fonctionnelle (Calvert *et al.*, 1997; van Atteveldt *et al.*, 2004; Pekkola *et al.*, 2005) qu'en magnétoencéphalographie (Levänen *et al.*, 1998; Lütkenhöner *et al.*, 2002; Gobbele *et al.*, 2003).

Chez le singe, des convergences multisensorielles dans le cortex auditif ont été confirmées par des études combinant des enregistrements intra-crâniens unitaires et multi-unitaires avec des mesures en potentiels évoqués (Schroeder *et al.*, 2001; Schroeder et Foxe, 2002; Brosch *et al.*, 2005; Ghazanfar *et al.*, 2005). Certaines études détaillées utilisent des enregistrements par microélectrodes chez le signe anesthésié, ce qui a permis de confirmer l'existence de réponses somatosensorielles au niveau neuronal dans des régions auditives directement adjacentes aux aires auditives primaires (Fu *et al.*, 2003). De façon similaire, la neuroimagerie par IRM fonctionnelle chez le singe anesthésié a démontré une suractivation audiovisuelle de la région auditive située postérieurement et latéralement par rapport au cortex auditif primaire (Kayser *et al.*, 2005). En plus des afférences visuelles et somatosensorielles, on a également trouvé des afférences concernant la position des yeux dans les aires auditives de bas niveau chez le singe, ces dernières incluant nettement les aires auditives primaires (Werner-Reiss *et al.*, 2003; Fu *et al.*, 2004).

Comme pour l'homme, la modulation de l'activité du cortex auditif primaire par des stimuli non-auditifs est particulièrement nette dans le cas de certains stimuli audiovisuels relatifs à la communication inter-individuelle chez le singe (Ghazanfar *et al.*, 2005). Dans ce cas, la quasi-totalité des réponses d'intégration multisensorielle observées dans les aires auditives primaires ou immédiatement adjacentes sont spécifiques à l'intégration audiovisuelle de la voix et du visage ; ainsi si le visage est remplacé par un stimulus visuel dynamique imitant les mouvements de la bouche, on observe une diminution en moyenne de l'intégration (Ghazanfar *et al.*, 2005). Cette intégration spécifique de la voix et du visage est également observée chez l'homme (Bernstein *et al.*, 2004; Schwartz *et al.*, 2004).

Traitement auditif et somatosensoriel dans les aires visuelles

Des études de cartographie à grande échelle chez l'animal ont permis de démontrer des activations des aires unimodales visuelles par des stimuli auditifs et somatosensoriels. Ainsi, dans le cortex visuel du chat, un peu moins de la moitié des neurones répondent à des stimuli auditifs, ces neurones présentant par ailleurs une forte sélectivité à la direction de la source auditive (Morrell, 1972). L'une des premières études dans ce domaine chez le singe, utilisant des mesures en potentiels évoqués chez le singe anesthésié, a montré que les aires activées par des clics auditifs s'étendaient sur une grande partie du gyrus temporal supérieur, de l'insula

et certaines parties du lobe pariétal (Pribram *et al.*, 1954). Une étude ultérieure en imagerie fonctionnelle par TEP chez le singe vigile a abouti à des résultats similaires, les stimuli auditifs activant de larges portions du cortex incluant les aires visuelles unisensorielles dans les lobes temporaux et pariétaux (Poremba *et al.*, 2003). Des enregistrements unitaires chez des singes effectuant une tâche de réponse audiovisuelle ont également montré des réponses auditives dans le cortex inféro-temporal (Gibson et Maunsell, 1997). Cet ensemble de données suggère que le cortex temporal ventral intervient dans la représentation mentale amodale des objets (Ghazanfar et Schroeder, 2006), ce qui semble cohérent avec une récente étude en IRM fonctionnelle montrant que des régions du gyrus fusiforme sont activées à la fois par les visages familiers et par les voix familières (von Kriegstein *et al.*, 2005).

Les preuves expérimentales les plus probantes de la convergence audiovisuelle dans le cortex visuel proviennent d'expériences de traçage anatomique chez le singe. Ainsi les aires visuelles primaires reçoivent des afférences provenant du cortex auditif et de la face supérieure du STS (Falchier *et al.*, 2002; Rockland et Ojima, 2003; Clavagnier *et al.*, 2004). Par ailleurs la répartition de ces afférences n'est pas uniforme, celles-ci étant plus spécifiquement concentrées dans la représentation du champ visuel périphérique (Falchier *et al.*, 2002) et dans les couches corticales 1 et 6, ce qui en fait des projections de type "feedback" (Rockland et Ojima, 2003). Ainsi le système auditif serait théoriquement en mesure d'influencer des phénomènes perceptifs visuels précoces tels qu'une simple détection de stimuli dans la partie périphérique du champ visuel, consacrée principalement à la détection de stimuli à faible contraste et non pas à leur identification. Des résultats similaires chez des souris anophtalmiques congénitales ont également montré la présence de connexions des aires auditives vers les aires visuelles primaires (Laemle *et al.*, 2006).

De nombreuses observations chez l'homme montrent que les aires corticales visuelles peuvent également être activées lors de la perception tactile (Sathian et Zangaladze, 2002). Ainsi la discrimination et l'identification haptiques d'objets activent le complexe latéral occipital, une région correpondant au cortex inféro-temporal chez le singe (Amedi *et al.*, 2001; James *et al.*, 2002; Pietrini *et al.*, 2004), tandis que la perception du mouvement tactile active le complexe MT (Hagen *et al.*, 2002; Blake *et al.*, 2004).

Traitement auditif et visuel dans les aires somatosensorielles

De façon générale, très peu d'études se sont intéressées à la présence de réponses multisensorielles dans les aires corticales somatosensorielles de bas niveau. Chez des singes entraînés à faire des associations visuo-haptiques ou audio-haptiques, un sous-ensemble des neurones corticaux somatosensoriels se met à répondre à la fois au stimulus tactile et aux indices visuels ou auditifs (Zhou et Fuster, 2000, 2004). Comme pour les autres types d'associations multisensorielles, les meilleures preuves de l'existence d'une convergence multisensorielle résident dans les études par traçage anatomique ; ainsi on a démontré chez le singe marmouset la présence de projections des aires visuelles FST et MT vers les aires somatosensorielles de bas niveau, ainsi que des connexions entre le cortex auditif et plusieurs aires somatosensorielles (Cappe et Barone, 2005).

5.4.3 Intégration dans les régions sous-corticales

De nombreuses observations ont révélé l'existence dans les étages sous-corticaux, en particulier dans le colliculus supérieur, de neurones multisensoriels pouvant constituer le substrat anatomique expliquant les processus d'intégration multimodale observés au niveau comportemental (Wallace *et al.*, 1996; Frens et Van Opstal, 1998; Bell *et al.*, 2001).

Les cartes spatiales visuelles et auditives coïncident en grande partie dans le colliculus supérieur, cette coïncidence audiovisuelle se retrouvant également au niveau de la cellule isolée (Gordon, 1975; Middlebrooks et Knudsen, 1984; King et Palmer, 1985; Meredith et Stein, 1986). Ainsi, des stimuli auditifs et visuels en coïncidence spatiale évoqueront des réponses d'une même cellule nerveuse à la fois dans son champ récepteur visuel et dans son champ récepteur auditif, ce qui induira une potentialisation mutuelle des réponses nerveuses.

Dans le colliculus inférieur du primate, une modification de la position oculaire module le niveau des réponses neuronales à un stimulus auditif, sans toutefois modifier la position de leurs champs récepteurs (Groh *et al.*, 2001). Au contraire, dans le colliculus supérieur du chat et du singe, les champs récepteurs auditifs se décalent en fonction de la position des yeux, de telle façon que ce décalage auditif compense en partie le décalage visuel provoqué par le changement de position des yeux (Jay et Sparks, 1984, 1987; Hartline *et al.*, 1995; Peck *et al.*, 1995).

Bibliographie

Alais, D. et Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.*, 14(3):257–262.

Amedi, A., Malach, R., Hendler, T., Peled, S. et Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.*, 4(3):324–330.

Andersen, R. A., Snyder, L. H., Bradley, D. C. et Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.*, 20:303–330.

Avillac, M., Deneve, S., Olivier, E., Pouget, A. et Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nat. Neurosci.*, 8(7):941–949.

Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W. et Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.*, 17(3):377–391.

Bell, A. H., Corneil, B. D., Meredith, M. A. et Munoz, D. P. (2001). The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. *Can. J. Exp. Psychol.*, 55(2):123–132.

Benevento, L. A., Fallon, J., Davis, B. J. et Rezak, M. (1977). Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Exp. Neurol.*, 57(3):849–872.

Bernstein, I. H., Clark, M. H. et Edelstein, B. A. (1969). Effects of an auditory signal on visual reaction time. *J. Exp. Psychol.*, 80(3):567–569.

Bernstein, L. E., Auer, Jr., E. T. et Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Comm.*, 44:5–18.

Bertelson, P. (1998). Starting from the ventriloquist : The perception of multimodal events. *In* Sabourin, M., Craik, F. I. et Robert, M., éditeurs : *Advances in Psychological Science. II. Biological and Cognitive Aspects*, pages 419–439. Hove, UK : Psychological Press.

Bertelson, P. et Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychon. Bull. Rev.*, 5:482–489.

Bertelson, P. et Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.*, 29(6):578–584.

Bertelson, P., Vroomen, J., de Gelder, B. et Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.*, 62(2):321–332.

Besle, J., Fort, A., Delpuech, C. et Giard, M.-H. (2004). Bimodal speech : early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.*, 20(8):2225–2234.

Blake, R., Sobel, K. V. et James, T. W. (2004). Neural synergy between kinetic vision and touch. *Psychol. Sci.*, 15(6):397–402.

Bremmer, F., Klam, F., Duhamel, J. R., Ben Hamed, S. et Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur. J. Neurosci.*, 16(8):1569–1586.

Brosch, M., Selezneva, E. et Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.*, 25(29):6797–6806.

Bruce, C., Desimone, R. et Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.*, 46(2):369–384.

Bushara, K. O., Grafman, J. et Hallett, M. (2001). Neural correlates of auditory-visual stimulus onset asynchrony detection. *J. Neurosci.*, 21(1):300–304.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D. et David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276:593–596.

Cappe, C. et Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur. J. Neurosci.*, 22:2886–2902.

Choe, C. S., Welch, R. B., Gilford, R. M. et Juola, J. F. (1975). The "ventriloquist effect" : visual dominance or response bias ? *Percept. Psychophys.*, 18:55–60.

Clavagnier, S., Falchier, A. et Kennedy, H. (2004). Long-distance feedback projections to area V1 : implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.*, 4(2):117–126.

Corneil, B. D. et Munoz, D. P. (1996). The influence of auditory and visual distractors on human orienting gaze shifts. *J. Neurosci.*, 16(24):8193–8207.

Corneil, B. D., Van Wanrooij, M., Munoz, D. P. et Van Opstal, A. J. (2002). Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J. Neurophysiol.*, 88(1):438–454.

Dixon, N. F. et Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, 9:719–721.

Dubose, C., Cardello, A. et Maller, O. (1980). Effects of colorants and flavorants on identification, perceived flavor intensity, and hedonic quality of fruit-flavored beverages and cake. *J. Food Sci.*, 45:1393–1418.

Duhamel, J. R., Colby, C. L. et Goldberg, M. E. (1998). Ventral intraparietal area of the macaque : congruent visual and somatic response properties. *J. Neurophysiol.*, 79(1):126–136.

Engelken, E. J. et Stevens, K. W. (1989). Saccadic eye movements in response to visual, auditory, and bisensory stimuli. *Aviat. Space Environ. Med.*, 60(8):762–768.

Falchier, A., Clavagnier, S., Barone, P. et Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.*, 22:5749–5759.

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M. et Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *J. Neurophysiol.*, 76(1):141–157.

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C. et Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by highdensity electrical mapping. *Brain Res. Cogn. Brain Res.*, 10:77–83.

Frassinetti, F., Bolognini, N. et Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.*, 147(3):332–343.

Frens, M. A. et Van Opstal, A. J. (1998). Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res. Bull.*, 46(3):211–224.

Frens, M. A., Van Opstal, A. J. et Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Percept. Psychophys.*, 57(6):802–816.

Fu, K. M., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., Garraghty, P. E. et Schroeder, C. E. (2003). Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.*, 23(20):7510–7515.

Fu, K. M., Shah, A. S., O'Connell, M. N., McGinnis, T., Eckholdt, H., Lakatos, P., Smiley, J. et Schroeder, C. E. (2004). Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. *J. Neurophysiol.*, 92(6):3522–3531.

Fuster, J. M., Bodner, M. et Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, 405(6784):347–351.

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L. et Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.*, 25(20):5004–5012.

Ghazanfar, A. A., Neuhoff, J. G. et Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proc. Natl. Acad. Sci. U.S.A.*, 99(24):15755–15757.

Ghazanfar, A. A. et Schroeder, C. E. (2006). Is neocortex essentially multisensory ? *Trends Cogn. Sci.*, 10(6):278–285.

Giard, M.-H. et Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans : A behavioral and electrophysiological study. *J. Cogn. Neurosci.*, 11:473–490.

Gibson, J. R. et Maunsell, J. H. (1997). Sensory modality specificity of neural activity related to memory in visual cortex. *J. Neurophysiol.*, 78(3):1263–1275.

Gielen, S. C., Schmidt, R. A. et Van den Heuvel, P. J. (1983). On the nature of intersensory facilitation of reaction time. *Percept. Psychophys.*, 34(2):161–168.

Giray, M. et Ulrich, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *J. Exp. Psychol. Hum. Percept. Perform.*, 19:1278–1291.

Gobbele, R., Schurmann, M., Forss, N., Juottonen, K., Buchner, H. et Hari, R. (2003). Activation of the human posterior parietal and temperopariatal cortices during audiotactile interaction. *Neuroimage*, 20:503–511.

Goldring, J. E., Dorris, M. C., Corneil, B. D., Ballantyne, P. A. et Munoz, D. P. (1996). Combined eye-head gaze shifts to visual and auditory targets in humans. *Exp. Brain Res.*, 111(1):68–78.

Gordon, B. G. (1975). Superior colliculus : structure, physiology and possible functions. *MTP Int. Rev. Sci.*, 3:185–230.

Graziano, M. S., Reiss, L. A. et Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397(6718):428–430.

Graziano, M. S., Taylor, C. S. et Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, 34(5):841–851.

Graziano, M. S., Yap, G. S. et Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science*, 266(5187):1054–1057.

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R. et Inati, S. (2001). Eye position influences auditory responses in primate inferior colliculus. *Neuron*, 29(2):509–518.

Hagen, M. C., Franzen, O., McGlone, F., Essick, G., Dancer, C. et Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *Eur. J. Neurosci.*, 16(5):957–964.

Harrington, L. K. et Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Exp. Brain Res.*, 122(2):247–252.

Hartline, P. H., Vimal, R. L., King, A. J., Kurylo, D. D. et Northmore, D. P. (1995). Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Exp. Brain Res.*, 104(3):402–408.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *J. Exp. Psychol.*, 63:289–293.

Hikosaka, K. (1993). The polysensory region in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Biomed. Res.*, 14:41–45.

Howard, I. P. et Templeton, W. B. (1966). *Human Spatial Orientation*. New York : Wiley.

Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G. et Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing : saccades versus manual responses. *J. Exp. Psychol. Hum. Percept. Perform.*, 20(1):131–153.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S. et Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, 40(10):1706–1714.

Jay, M. F. et Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, 309(5966):345–347.

Jay, M. F. et Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus. II. coordinates of auditory signals. *J. Neurophysiol.*, 57(1):35–55.

Jiang, W., Jiang, H. et Stein, B. E. (2002). Two corticotectal areas facilitate multisensory orientation behavior. *J. Cogn. Neurosci.*, 14(8):1240–1255.

Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G. et Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol.*, 78(6):2834–2847.

Kayser, C., Petkov, C. I., Augath, M. et Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron*, 48(2):373–384.

King, A. J. et Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.*, 60(3):492–500.

Laemle, L. K., Strominger, N. L. et Carpenter, D. O. (2006). Cross-modal innervation of primary visual cortex by auditory fibers in congenitally anophthalmic mice. *Neurosci. Lett.*, 396(2):108–112.

Leinonen, L., Hyvarinen, J. et Sovijarvi, A. R. (1980). Functional properties of neurons in the temporo-parietal association cortex of awake monkey. *Exp. Brain Res.*, 39(2):203–215.

Levänen, S., Jousmaki, V. et Hari, R. (1998). Vibration-induced auditory cortex activation in a congenitally deaf adult. *Curr. Biol.*, 8:869–872.

Lewkowicz, D. J. (1996). Perception of auditory-visual temporal synchrony in human infants. *J. Exp. Psychol. Hum. Percept. Perform.*, 22:1094–1106.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception :an epigenetic systems/limitations view. *Psychol. Bull.*, 126:281–308.

Lütkenhöner, B., Lammertmann, C., Simoes, C. et Hari, R. (2002). Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage*, 15:509–522.

Maier, J. X., Neuhoff, J. G., Logothetis, N. K. et Ghazanfar, A. A. (2004). Multisensory integration of looming signals by rhesus monkeys. *Neuron*, 43(2):177–181.

Mazzoni, P., Bracewell, R. M., Barash, S. et Andersen, R. A. (1996). Spatially tuned auditory responses in area LIP of macaques performing delayed memory saccades to acoustic targets. *J. Neurophysiol.*, 75(3):1233–1241.

McDonald, J. J., Teder-Salejarvi, W. A. et Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, 407(6806):906–908.

Meredith, M. A. et Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.*, 56(3):640–662.

Meredith, M. A. et Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J. Neurophysiol.*, 75(5):1843–1857.

Middlebrooks, J. C. et Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.*, 4(10):2621–2634.

Miller, J. O. (1986). Time course of coactivation in bimodal divided attention. *Percept. Psychophys.*, 40(5):331–343.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E. et Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans : A high-density electrical mapping study. *Brain Res. Cogn. Brain Res.*, 14:115–128.

Morrell, F. (1972). Visual system's view of acoustic space. *Nature*, 238(5358):44–46.

Morrot, G., Brochet, F. et Dubourdieu, D. (2001). The color of odors. *Brain Lang.*, 79(2):309–320.

Murray, M. M., Molholm, S., Michel, C. M., Ritter, W., Heslenfeld, D. J., Schroeder, C. E., Javitt, D. C. et Foxe, J. J. (2004). Grabbing your ear : rapid auditory-somatosensory multisensory interactions in lowlevel sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex*, 15:963–974.

Peck, C. K., Baro, J. A. et Warder, S. M. (1995). Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus. *Exp. Brain Res.*, 103(2):227–242.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A. et Sams, M. (2005). Primary auditory cortex activation by visual speech : An fMRI study at 3T. *Neuroreport*, 16:125–128.

Perrott, D. R., Saberi, K., Brown, K. et Strybel, T. Z. (1990). Auditory psychomotor coordination and visual search performance. *Percept. Psychophys.*, 48(3):214–226.

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W. H., Cohen, L., Guazzelli, M. et Haxby, J. V. (2004). Beyond sensory images : object-based representation in the human ventral pathway. *Proc. Natl. Acad. Sci. U.S.A.*, 101(15):5658–5663.

Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L. et Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science*, 299(5606):568–572.

Pribram, K. H., Rosner, B. S. et Rosenblith, W. A. (1954). Electrical responses to acoustic clicks in monkey : extent of neocortex activated. *J. Neurophysiol.*, 17(4):336–344.

Radeau, M. et Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs. thomas (1941) revisited. *Psychol. Res.*, 49(1):17–22.

Rockland, K. S. et Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.*, 50(1-2):19–26.

Sathian, K. et Zangaladze, A. (2002). Feeling with the mind's eye : contribution of visual cortex to tactile perception. *Behav. Brain Res.*, 135(1-2):127–132.

Schiff, W., Caviness, J. A. et Gibson, J. J. (1962). Persistent fear responses in rhesus monkeys to the optical stimulus of "looming". *Science*, 136:982–983.

Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K. P. et Bremmer, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *J. Neurosci.*, 25(18):4616–4625.

Schroeder, C. E. et Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.*, 14(1):187–198.

Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, J. F. et Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.*, 85(3):1322–1327.

Schwartz, J.-L., Berthommier, F. et Savariaux, C. (2004). Seeing to hear better : evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2):B69–B78.

Sekuler, R., Sekuler, A. B. et Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385(6614):308.

Shams, L., Kamitani, Y. et Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(6814):788.

Shams, L., Ma, W. J. et Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport*, 16(17):1923–1927.

Slutsky, D. A. et Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, 12(1):7–10.

Stein, B. E., Huneycutt, W. S. et Meredith, M. A. (1988). Neurons and behavior : the same rules of multisensory integration apply. *Brain Res.*, 448(2):355–358.

Stein, B. E. et Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA : MIT Press.

Stein, B. E., Meredith, M. A., Huneycutt, W. S. et McDade, L. (1989). Behavioral indices of multisensory integration : orientation to visual cues is affected by auditory stimuli. *J. Cogn. Neurosci.*, 1:12.

van Atteveldt, N., Formisano, E., Goebel, R. et Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, 43(2):271–282.

van Wassenhove, V., Grant, K. W. et Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.*, 102(4):1181–1186.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P. et Giraud, A.-L. (2005). Interaction of face and voice areas during speaker recognition. *J. Cogn. Neurosci.*, 17(3):367–376.

Vroomen, J., Bertelson, P. et de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Percept. Psychophys.*, 63(4):651–659.

Wallace, M. T., Wilkinson, L. K. et Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *J. Neurophysiol.*, 76(2):1246–1266.

Watanabe, K. et Shimojo, S. (1998). Attentional modulation in perception of visual motion events. *Perception*, 27(9):1041–1054.

Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M. et Groh, J. M. (2003). Eye position affects activity in primary auditory cortex of primates. *Curr. Biol.*, 13(7):554–562.

Wilkinson, L. K., Meredith, M. A. et Stein, B. E. (1996). The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Exp. Brain Res.*, 112(1):1–10.

Witkin, H. A., Wapner, S. et Leventhal, T. (1952). Sound localization with conflicting visual and auditory cues. *J. Exp. Psychol.*, 43(1):58–67.

Zahn, J. R., Abel, L. A. et Dell'Osso, L. F. (1978). Audio-ocular response characteristics. *Sens. Processes*, 2(1):32–37.

Zhou, Y. D. et Fuster, J. M. (2000). Visuo-tactile cross-modal associations in cortical somatosensory cells. *Proc. Natl. Acad. Sci. U.S.A.*, 97:9777–9782.

Zhou, Y. D. et Fuster, J. M. (2004). Somatosensory cell responses to an auditory cue in a haptic memory task. *Behav. Brain Res.*, 153:573–578.

Deuxième partie

Surdité, implant cochléaire et réorganisation corticale

Chapitre 6

La surdité

Sommaire

6.1 Qu'est-ce que la surdité ? . . . 128
6.1.1 Classification des surdités . . . 128
6.1.2 Prévalence de la surdité dans le monde . . . 129
6.2 Dépistage de la surdité . . . 129
6.2.1 Symptômes . . . 129
6.2.2 Acoumétrie . . . 130
6.2.3 Audiométrie . . . 130
6.2.4 Examens complémentaires . . . 132
6.3 Causes de la surdité . . . 134
6.3.1 Etiologie des surdités de transmission . . . 134
6.3.2 Etiologie des surdités de perception . . . 135
6.4 Effets de la surdité sur la perception de la parole . . . 137
6.4.1 Perception auditive de la parole lors de la surdité . . . 137
6.4.2 Perception visuelle de la parole lors de la surdité . . . 137
6.4.3 Perception audiovisuelle de la parole lors de la surdité . . . 138
6.5 Prothèses auditives . . . 138
6.5.1 Prothèses ossiculaires . . . 138
6.5.2 Aides auditives externes . . . 138
6.5.3 Implants d'oreille moyenne . . . 139
6.5.4 Neuroprothèses auditives . . . 139
6.5.5 Prothèses à ancrage osseux . . . 139
6.6 Communication avec les personnes sourdes . . . 140
6.6.1 La langue des signes . . . 140
6.6.2 Méthodes oralistes de communication . . . 140
Bibliographie . . . 143

Introduction

La surdité touche l'être humain dans sa capacité à utiliser le langage verbal pour communiquer avec ses semblables, ce qui est non seulement une spécificité particulière à l'homme mais également une fonction biologique indispensable à une bonne structuration

psychologique et sociale. Selon les sources statistiques, entre 10 et 20 % de la population mondiale est affectée à des degrés divers par la surdité, cette proportion étant moindre pour les jeunes et considérablement plus importante pour les personnes âgées. Nous aborderons dans ce chapitre quelques notions générales relatives à la surdité, en préliminaire aux chapitres suivants, consacrés à l'implant cochléaire ainsi qu'aux aspects neurofonctionnels de la surdité et de l'implantation cochléaire.

6.1 Qu'est-ce que la surdité ?

6.1.1 Classification des surdités[1]

On appelle déficience auditive, ou surdité, toute baisse unilatérale ou bilatérale de l'ouïe, quels que soient son degré et sa localisation.

En fonction du moment d'apparition de la surdité, on parle de surdité congénitale† lorsque les personnes sont nées sourdes, de surdité prélinguale† lorsque l'audition est affectée lors de la petite enfance, avant l'acquisition du langage parlé, et de surdité postlinguale† lorsque l'audition est affectée après la période critique d'acquisition du langage.

En fonction du point du système auditif qui est touché, on parle de surdité de transmission† s'il y a atteinte des organes intervenant dans la transmission physique des sons (comme le conduit auditif externe, le tympan ou les osselets), de surdité de perception† s'il y a atteinte des organes intervenant dans le codage ou la transmission nerveuse des sons (comme la cochlée ou le nerf auditif), et de surdité centrale† s'il y a atteinte des centres nerveux auditifs.

Enfin, en fonction de l'évolution temporelle de la surdité, on fait une distinction entre la surdité progressive†, qui évolue lentement avec une dégradation progressive de l'audition, et la surdité brusque†, qui apparaît au cours d'un laps de temps très restreint, entre quelques minutes et plusieurs jours.

De façon générale, les déficiences auditives sont associées à une perte de la perception des sons, de la parole en particulier, qui comporte des sons aigus et des sons graves dont la puissance acoustique est variable, et ne peut être réduite à un niveau acoustique moyen. Après un bilan clinique, la mesure audiométrique est réalisée dans des conditions acoustiques satisfaisantes. Elle fait apparaître une perte en décibels par rapport à l'oreille normale en référence aux normes ISO. Une perte totale moyenne est calculée à partir de la perte en dB aux fréquences 500 Hz, 1000 Hz, 2000 Hz et 4000 Hz. Toute fréquence non perçue est notée à 120 dB de perte. Leur somme est divisée par quatre arrondie à l'unité supérieure. En cas de surdité asymétrique, le niveau moyen de perte en dB est multiplié par 7 pour la meilleure oreille et par 3 pour la plus mauvaise oreille. La somme est divisée par 10.

Audition normale ou subnormale La perte tonale moyenne ne dépasse pas 20 dB. Il s'agit éventuellement d'une atteinte tonale légère sans incidence sociale.

Déficience auditive légère La perte tonale moyenne est comprise entre 21 dB et 40 dB. La parole est perçue à voix normale, elle est difficilement perçue à voix basse ou lointaine. La plupart des bruits familiaux sont perçus.

Déficience auditive moyenne *Premier degré* : la perte tonale moyenne est comprise entre 41 et 55 dB. *Deuxième degré* : la perte tonale moyenne est comprise entre 56 et 70 dB. La parole est perçue si on élève la voix. Le sujet comprend mieux en regardant parler. Quelques bruits familiers sont encore perçus.

[1] d'après BIAP (1997)

Déficience auditive sévère *Premier degré* : la perte tonale moyenne est comprise entre 71 et 80 dB. *Deuxième degré* : la perte tonale moyenne est comprise entre 81 et 90 dB. La parole est perçue à voix forte près de l'oreille. Les bruits forts sont perçus.

Déficience auditive profonde *Premier degré* : la perte tonale moyenne est comprise entre 91 et 100 dB. *Deuxième degré* : la perte tonale moyenne est comprise entre 101 et 110 dB. *Troisième degré* : la perte tonale moyenne est comprise entre 111 et 119 dB. Aucune perception de la parole. Seuls les bruits très puissants sont perçus.

Déficience auditive totale - Cophose La perte moyenne est de 120 dB. Rien n'est perçu.

6.1.2 Prévalence de la surdité dans le monde

On dispose de peu de données statistiques concernant la surdité dans les différents pays du monde. Les quelques études menées dans les pays européens montrent que, selon les pays, entre 5 et 15% des habitants souffrent de déficiences auditives. Pour la population française, une évaluation du Ministère de la santé indique une prévalence de l'ordre de 6.6% en 2005, répartie de la façon suivante :

- 3.72 % de déficiences auditives légères ;
- 2.10 % de déficiences auditives moyennes ;
- 0.60 % de déficiences auditives sévères ;
- 0.18 % de déficiences auditives profondes.

La prévalence des déficiences auditives augmente de façon importante avec l'âge. Il faut par ailleurs noter que la surdité est devenue en 1982 la première cause de maladie professionnelle en France. La surdité constitue donc un problème de santé publique de grande ampleur.

6.2 Dépistage de la surdité

De nos jours le diagnostic de la surdité passe par une série d'étapes clairement codifiées, depuis le simple interrogatoire diagnostique du patient, suivi par les examens classiques d'acoumétrie et d'audiométrie, jusqu'à l'emploi de techniques sophistiquées d'investigation des propriétés physiologiques, électrophysiologiques et anatomiques du système auditif.

On dispose à l'heure actuelle d'un éventail très large de techniques de dépistage et de diagnostic de la surdité, aussi bien chez le nouveau-né et le jeune enfant que chez l'adulte. En ce qui concerne les nouveaux-nés, le contrôle médical de l'audition a été introduit dans plusieurs pays afin de détecter tôt les déficiences auditives (O'Donoghue, 1996; O'Donoghue *et al.*, 1998). En effet l'établissement du diagnostic de la surdité congénitale, dont l'incidence est d'environ un cas sur mille, est rendu difficile par le fait que la détection d'une perte auditive n'est possible qu'après la naissance, tandis que la cochlée humaine devient fonctionnelle dès la vie intra-utérine (Granier-Deferre *et al.*, 1985). Cela explique que les cliniciens ne font de différence qu'entre la surdité prélinguale (acquise avant le développement du langage) et la surdité postlinguale (acquise après le développement du langage).

6.2.1 Symptômes

L'apparition d'une surdité chez l'adulte doit entraîner un certain nombre d'investigations permettant de savoir à quel niveau se pose le problème et quelle en est la cause, de préciser le degré de perte auditive ainsi que son retentissement sur l'intelligibilité, puis enfin d'adapter le traitement au cas individuel du patient.

Le premier interrogatoire diagnostique du patient permet de lui faire préciser les élements suivants :

- la date et les circonstances d'apparition de l'hypoacousie (traumatisme crânien, virose, prise médicamenteuse, etc.) ;
- les signes d'accompagnement : acouphènes (bourdonnements d'oreille), otorrhée (écoulement de liquide par l'oreille), otorragie (écoulement de sang par l'oreille), otalgie (douleur dans l'oreille), vertiges, paralysie faciale, céphalées ;
- l'allure évolutive de la surdité : progressive, fluctuante ou brutale ;
- le caractère uni- ou bilatéral de la surdité ;
- le retentissement socio-professionnel de la perte auditive ;
- la profession ;
- les antécédents oto-rhino-laryngologiques ;
- les antécédents familiaux de surdité.

L'examen du tympan à l'otoscope est fondamental car il permet de distinguer les surdités à tympan normal et les surdités à tympan anormal (inflammatoire, perforé, etc.). Il permet également d'éliminer un bouchon de cérumen ou une otite externe. Les autres examens sont demandés en fonction du contexte. Pour les surdités de transmission, ils peuvent inclure une impédancemétrie ainsi qu'une radio ou un scanner des rochers, tandis que pour les surdités de perception, ils peuvent inclure une étude des potentiels évoqués auditifs, une électrocochléographie, un test au glycérol, une électronystagmographie ainsi qu'un scanner par tomodensiométrie ou par résonance magnétique nucléaire.

6.2.2 Acoumétrie

L'acoumétrie fait partie des toutes premières étapes du diagnostic de la surdité. Simple et rapide, elle se fait au moyen d'un diapason et permet de distinguer les surdités de transmission et de perception.

Test de Weber

Le test de Weber[†] consiste à comparer l'acuité auditive des deux oreilles à l'aide d'un diapason appuyé sur le front. Le médecin demande au patient de quel côté il perçoit le son. Si le son est perçu par l'oreille la plus sourde, il s'agit d'une surdité de transmission ; s'il est mieux perçu par l'oreille saine, il s'agit d'une surdité de perception (lésions de l'oreille interne). En cas d'audition normale ou symétrique, les vibrations sont entendues des deux côtés.

Test de Rinne

Le test de Rinne[†] sert à diagnostiquer les surdités de transmission. Un diapason est placé devant l'oreille, puis contre la mastoïde. A l'état normal, le son est mieux et plus longtemps perçu par voie aérienne. Dans les maladies de l'appareil de transmission, il est mieux perçu par voie osseuse : le Rinne est alors négatif.

6.2.3 Audiométrie

Audiométrie tonale liminaire

L'audiométrie tonale liminaire consiste à évaluer les niveaux seuils de l'audition pour chacune des deux oreilles sur l'éventail des sept principales fréquences du spectre auditif de l'homme : 125 et 250 Hz (basses fréquences), 500, 1000 et 2000 Hz (fréquences moyennes), 4000 et 8000

Hz (hautes fréquences). L'évaluation des seuils se fait en employant des sons purs émis à différentes intensités sur toute l'échelle des niveaux physiologiques, du seuil d'audition moyen au seuil intolérable moyen, ce qui permet de mesurer les seuils perceptifs du sujet en fonction de la fréquence. On trouvera plus de détails sur les limites perceptives du système auditif humain dans « *Perception auditive* », § 1.2, p. 26.

On explore ainsi deux manières d'entendre. L'audition par voie aérienne est testée à l'aide d'un haut-parleur émettant des sons en champ libre (condition binaurale) ou avec un casque muni d'écouteurs (condition monaurale). Ce procédé permet d'évaluer l'audition dans des conditions physiologiques normales. L'audition par voie osseuse est testée grâce à un vibreur placé sur l'os de la mastoïde, derrière l'oreille. Les vibrations émises se transmettent par conduction solidienne à toute la boîte crânienne et parviennent à la cochlée, où elles font entrer en vibration les liquides cochléaires et la membrane basilaire. Ce procédé permet de court-circuiter l'action de l'oreille externe et de l'oreille moyenne, ce qui permet d'évaluer correctement la contribution perceptive de l'oreille interne, du nerf et des centres cérébraux auditifs (pour les structures anatomiques de l'audition, voir « *Système auditif périphérique* », sect. 1.1, p. 18). De cette façon, on peut faire la distinction entre une surdité de transmission et une surdité de perception.

Les résultats audiométriques sont reportés sur l'audiogramme, qui comporte l'échelle des fréquences en abscisses et l'échelle des niveaux acoustiques en ordonnées négatives. Les niveaux perceptifs sont appréciés en fonction de leur écart aux seuils physiologiques normaux ; ils sont mesurés en dB HL (*hearing level*). Une audition normale aboutit à un audiogramme pour lequel les deux courbes de conduction aérienne et de conduction osseuse sont confondues avec la ligne à 0 dB HL, qui correspond au seuil liminaire de l'audition normale.

On peut également utiliser l'audiométrie tonale liminaire conditionnée, qui consiste à développer un réflexe conditionné lors de la perception d'un son (comme appuyer sur un bouton pour faire apparaître une image) ; elle est utilisée pour l'examen des enfants.

Audiométrie tonale supraliminaire, ou test de Fowler

L'audiométrie tonale supraliminaire, ou test de Fowler, consiste à évaluer pour chacune des oreilles le niveau sonore perçu subjectivement par la personne. Pour chaque fréquence, on va donc rechercher les niveaux d'intensité sonore pour l'oreille gauche et l'oreille droite aboutissant au même niveau sonore perçu. Les résultats sont typiquement portés sur un graphique qui exprime les intensités sonores pour l'oreille droite en fonction des intensités pour l'oreille gauche.

Le test de Fowler permet de rechercher un phénomène de recrutement, qui correspond à une diminution de la zone d'audition par compression de l'échelle des sensations d'intensité sonore. En cas de recrutement simple, seul le seuil absolu est plus élevé, ce qui se traduit par des difficultés pour entendre les sons d'intensité faibles. En cas de sur-recrutement, non seulement le seuil absolu est plus élevé (comme dans le cas du recrutement simple) mais le seuil douloureux est également abaissé, ce qui se traduit par une hypersensibilité aux intensités fortes.

Le phénomène de recrutement n'existe que si le nerf est intact et témoigne d'une lésion limitée à l'organe de Corti. Dans le cas d'une surdité de perception, les potentiels nerveux des cellules lésées sont trop faibles, ce qui amène les fibres nerveuses à « recruter » des cellules supplémentaires pour permettre l'audition des faibles puissances. Ainsi, pour les puissances élevées, le signal devient trop fort et le patient a une sensation douloureuse. L'audiométrie supraliminaire permet de dissocier les surdités cochléaires (où il y a recrutement) des surdités rétro-cochléaires, le phénomène de recrutement n'étant pas observable en cas d'atteinte du nerf auditif.

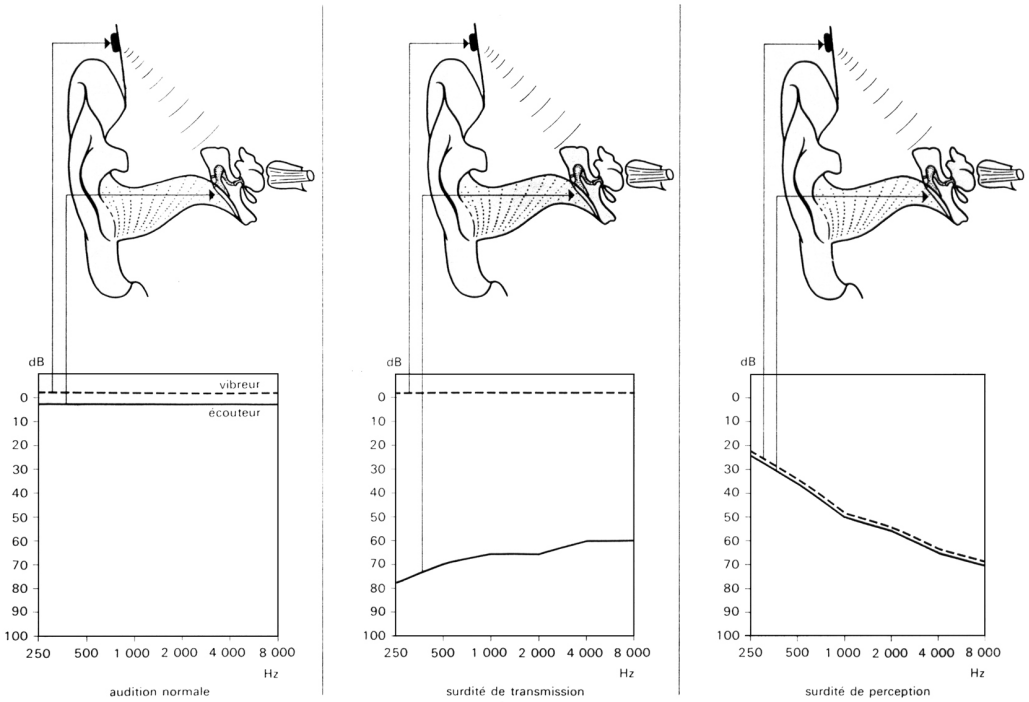


FIG. 6.1 – Analyse des résultats de l'examen audiométrique

Audiométrie vocale

L'audiométrie vocale utilise le stimulus le plus physiologique de l'ouïe, c'est-à-dire la voix humaine. Elle est utilisée principalement pour l'indication d'appareillage en cas de surdité. Elle consiste à étudier la perception par le patient de séries de mots courants (monosyllabiques, bisyllabiques ou plus complexes) émis pour différentes intensités sonores de prononciation : voix pianissimo ($\simeq$40 dB), voix mezzoforte ($\simeq$55 dB) et voix fortissimo ($\simeq$80 dB). On note alors le pourcentage de mots compris par le patient pour chaque niveau de puissance.

6.2.4 Examens complémentaires

Impédancemétrie

L'impédancemétrie auditive est une méthode qui permet d'explorer objectivement le fonctionnement de l'appareil de transmission de l'audition. Elle comporte deux examens : la tympanométrie et le réflexe stapédien.

La tympanométrie mesure l'impédance de l'ensemble formé par le tympan et la chaîne des osselets. Celle-ci est fonction de la différence de pression entre l'oreille moyenne et l'oreille externe (air ambiant dans le conduit auditif externe). En audiométrie, l'impédance est une grandeur permettant de caractériser la résistance d'une structure à la transmission d'une vibration acoustique.

Le réflexe stapédien fait intervenir la contraction du muscle fixé sur l'étrier. L'étrier est l'un des trois osselets de l'oreille moyenne qui permettent la transmission des vibrations captées par le tympan à l'oreille interne. La contraction réflexe de ce muscle est déclenchée par un bruit très intense. Ce réflexe permet d'autre part de protéger l'oreille contre le risque de saturation et

éventuellement le risque de survenue de lésions quand le bruit est particulièrement important. Il est à l'origine d'une modification de l'impédancemétrie acoustique qui est abolie quand un patient présente une surdité de transmission.

Potentiels évoqués auditifs

Les potentiels évoqués auditifs (PEA) permettent l'étude des fibres nerveuses des voies auditives situées entre la cochlée et les premières projections cérébrales (voir *« Système auditif central »*, sect. 1.3, p. 29). L'enregistrement des potentiels évoqués auditifs (PEA) permet d'obtenir un tracé fait de 7 ondes qui reflètent l'activité électrique du système nerveux auditif, chaque onde correspondant à une structure anatomique précise.

L'étude des potentiels évoqués auditifs se pratique dans une pièce isolée électriquement et insonorisée. Lors d'une séance de PEA, des sons brefs et intenses sont délivrés à travers un casque d'écoute. Les potentiels électriques sont recueillis au niveau de la peau située derrière les oreilles, à l'aide d'électrodes surfaciques. Le patient est allongé, de manière à relâcher la musculature du cou. Une électrode est placée au niveau du lobule de l'oreille testée, une autre sur le sommet du crâne et une troisième au niveau du lobule de l'oreille de l'autre côté. Chaque oreille est ensuite stimulée au casque par des clics à la cadence de 20 par seconde (il faut 2048 stimulations pour obtenir un tracé interprétable), en commençant par de fortes intensités (100 dB) suivies d'une diminution progressive jusqu'à l'obtention du seuil physiologique. L'examen dure 30 minutes par oreille.

Les PEA permettent en particulier d'évaluer les réponses du tronc cérébral à une stimulation auditive. Ils fournissent un moyen de diagnostiquer de façon objective l'état de préservation du système auditif périphérique, en mesurant directement les réponses nerveuses à un signal sonore. Lorsque l'intensité de la stimulation sonore augmente, on peut normalement observer un phénomène de recutement de latence, correspondant à une diminuation de latence de la réponse nerveuse, ainsi qu'un recrutement d'amplitude, correspondant à une augmentation d'amplitude de la réponse nerveuse.

Cette méthode peut en particulier être utilisée sur des patients pour lesquels l'audiogramme classique est irréalisable, par exemple pour de très jeunes enfants, pour des patients ne comprenant pas les directives du médecin ou pour des patients inconscients. Cet examen peut se faire dès la naissance pendant le sommeil physiologique après la tétée. Après 6 mois, une prémédication peut être nécessaire. A partir de 4 ans, l'examen ne nécessite aucune prémédication.

Electrocochléographie

L'électrocochléographie est un examen pratiqué sous anesthésie générale, consistant à enregistrer les réponses de la cochlée induites par des stimulations sonores très brèves, répétitives et de phase alternativement opposée. Il consiste à placer, sous vision microscopique, une très fine électrode à travers le tympan pour enregistrer l'activité électrique du nerf auditif. De cette façon, il est possible d'étudier la réponse à une stimulation électrique de la cochlée (notamment des cellules de Corti qui sont les récepteurs l'oreille interne) et du nerf auditif.

Electronystagmographie

L'électronystagmographie est un test qui permet d'évaluer l'intégrité du système vestibulaire (constitué des canaux semi-circulaires de l'oreille interne, du nerf vestibulaire, du tronc cérébral et du cervelet). Suite à des stimulations affectant le sens de l'équilibre, ce test permet de déclencher des mouvements de va-et-vient des yeux (nystagmus) intervenant dans le contrôle de

l'équilibre. On enregistre l'activité électrique associée à ces mouvements oculaires qui dépendent, entre autres, de cellules nerveuses du tronc cérébral et du cervelet.

Au cours de l'examen, le patient est assis sur une chaise pivotante dont la rotation permet de stimuler le système vestibulaire. On enregistre alors les mouvements oculaires correspondants grâce à des électrodes placées sur la peau autour des yeux. Certaines affections neurodégénératives, comme la sclérose en plaques, peuvent provoquer des troubles oculomoteurs. On observe fréquemment un nystagmus (secousses oculaires rapides dans une direction suivies d'un retour plus lent à la normale), ainsi que des problèmes d'ophtalmoplégie internucléaire.

6.3 Causes de la surdité

6.3.1 Etiologie des surdités de transmission

La surdité de transmission est liée à une lésion de l'oreille externe ou de l'oreille moyenne. Elle touche essentiellement la conduction aérienne et est caractérisée par une conduction osseuse normale (test de Rinne négatif et test de Weber latéralisé du côté atteint) et par une absence de recrutement lors du test de Fowler. Elle se traduit par des déficiences auditives légères ou moyennes. Chez l'enfant la cause la plus fréquente est l'otite séreuse, tandis que chez l'adulte la surdité peut être liée à des causes plus variées : otospongiose, agénésie, bouchon de cérumen, traumatisme, perforation du tympan, otite, etc.

Surdités de transmission sans atteinte du tympan

Dans le cas où l'examen otologique indique un tympan normal, la cause la plus fréquente de surdité de transmission est l'otospongiose[†], ou otosclérose[†]. Celle-ci se traduit par une évolution pathologique des tissus osseux aboutissant à la fixation de la platine de l'étrier dans la fenêtre ovale (otospongiose de premier type). Les principales autres causes de surdité de transmission à tympan normal sont les bouchons de cérumen obstruant le conduit auditif externe, les ruptures de la chaîne des osselets à la suite d'un traumatisme crânien, ainsi que les malformations ossiculaires.

Surdités de transmission avec atteinte du tympan

Si l'examen otologique indique une affection du tympan, on peut distinguer trois principales causes parmi les surdités de transmissions : les affections infectieuses ou inflammatoires (principalement représentées par les otites), les traumatismes mécaniques, et les tumeurs ou malformations affectant l'oreille moyenne ou externe.

Otites. Les otites chroniques et leurs séquelles représentent la cause majeure des surdités de transmission avec atteinte du tympan. Elles se déclinent sous différentes formes. Dans le cas d'une otite chronique simple non évolutive, l'oreille n'est plus affectée par aucun écoulement, mais il persiste une surdité de transmission en lien avec une perforation du tympan et une destruction plus ou moins importante des osselets. Dans le cas d'une otite chronique évolutive muqueuse simple, on peut observer une otorrhée récidivante, souvent associée à une infection chronique des voies aériennes supérieures.

Dans certains cas rares, l'otite chronique peut être accompagnée d'une complication avec tympanosclérose, qui résulte d'une calcification de la muqueuse tapissant la paroi interne du tympan, ce qui provoque un épaississement et une rigidification de ce dernier. En cas d'otite fibro-adhésive, la muqueuse subit une hypertrophie fibreuse susceptible de combler partiellement ou totalement la cavité de l'oreille moyenne. Enfin, la surdité de transmission peut également

être liée à une otite moyenne aigüe, à une otite séreuse (en particulier chez les enfants) ou encore à un catarrhe tubaire aigu (hypersécrétion de la trompe d'Eustache responsable d'une mauvaise aération de la caisse du tympan).

Traumatismes mécaniques. Des corps étrangers introduits dans le conduit auditif externe peuvent parfois provoquer une perforation du tympan voire une luxation des osselets, menant à une surdité de transmission. L'extraction de l'objet doit être pratiquée par un médecin ORL sous microscope pour éviter d'aggraver les lésions en refoulant le corps étranger dans la caisse. En cas de perforation tympanique, un traitement antibiotique est nécessaire. La présence de vertiges ou d'écoulements impose une exploration chirurgicale d'urgence car il y a risque de fistule périlymphatique avec évolution vers la surdité.

La surdité de transmission peut également être provoquée par certains traumatismes crâniens affectant les os du rocher, ainsi que par l'exposition à de trop grandes différences de pression de part et d'autre du tympan (barotraumatismes). Ces derniers peuvent typiquement se produire lors de plongées sous-marines ou de voyages en avion

Tumeurs et malformations. Parmi les possibles causes de surdité de transmission, il faut mentionner le cholestéatome, une tumeur bénigne provoquant un kyste envahissant à des degrés variables la caisse du tympan, l'attique et la mastoïde. Cliniquement, le cholestéatome est diagnostiqué en cas d'écoulement chronique fétide de l'oreille, rebelle aux traitements antibiotiques. Le scanner des rochers fait partie du bilan préopératoire. Le traitement chirurgical est obligatoire car la tumeur se propage aux tissus osseux avoisinants et peut provoquer des complications graves : paralysie faciale, labyrinthite avec cophose (surdité totale), méningite, abcès cérébral, etc. Certaines tumeurs ou malformations congénitales, bien que rares, peuvent également provoquer des surdités de transmission avec atteinte du tympan.

6.3.2 Etiologie des surdités de perception

Les surdités de perception sont des surdités neuro-sensorielles caractérisées par une atteinte de la cochlée ou des voies nerveuses situées en aval de la cochlée. Ces surdités sont la plupart du temps caractérisées par des difficultés de compréhension dans le bruit ambiant (les patients « entendent bien mais comprennent mal ») ainsi qu'une perception diminuée de la voix chuchotée et de la voix haute, et sont fréquemment accompagnées d'acouphènes.

L'examen otoscopique est normal et associé à un déficit auditif atteignant à la fois la conduction aérienne et osseuse, et ce plus particulièrement dans les aigüs. Le test de Weber est latéralisé du côté le plus sourd et le test de Rinne est positif. Le diagnostic est le plus souvent confirmé par l'audiométrie et l'examen des potentiels évoqués auditifs, bien que certains cas requièrent des tests supplémentaires comme l'électronystagmographie, les tests osmotiques, l'électrocochléographie, la recherche d'otoémissions provoquées ou un examen radioanatomique approfondi.

Surdités endocochléaires

Les surdités endocochléaires sont des surdités de perception caractérisées par une atteinte des cellules sensorielles de l'organe de Corti. Les étiologies possibles sont très diversifiées et peuvent être liées à la sénescence, à des traumatismes mécaniques ou encore à des causes génétiques, congénitales, virales, infectieuses, inflammatoires, toxiques (ototoxicité médicamenteuse), dégénératives (otospongiose cochléaire et maladies auto-immunes) ou physiopathologiques (maladie de Ménière).

Vieillissement. La presbyacousie est à l'audition ce que la presbytie est à la vision : elle traduit une détérioration lente et progressive de l'audition liée au vieillissement physiologique naturel de l'organe de Corti. Plus ou moins marquée selon les sujets, cette affection débute usuellement entre 45 et 55 ans et atteint en priorité la perception des fréquences aiguës. En effet les cellules ciliées associées aux hautes fréquences, situées à l'entrée de la cochlée, sont les plus fragiles et sont les premières à être détruites au cours de la vie. Ce type de maladie dû à l'âge est très fréquent et des appareils d'amplification sont généralement suffisants pour continuer à vivre normalement.

Traumatismes. Un traumatisme sonore peut être à l'origine d'une surdité de perception endocochléaire. Dans cette catégorie, le bruit présent sur le lieu de travail est souvent une cause de diminution des performances auditives. Si le niveau sonore dépasse rarement 40 dB dans les bureaux, il devient un réel problème dans l'industrie mécanique (marteaux perforateurs, machines à air comprimé, etc.), l'industrie du bois (scies circulaires), les métiers du tissage, de l'édition (rotatives), ainsi que les métiers associés au secteur du Bâtiment et travaux publics (marteaux-piqueurs). Dans ces derniers cas, la majorité des machines utilisées atteignent souvent un niveau sonore de 100 dB.

Les traumatismes crâniens peuvent également provoquer des dégâts importants dans l'oreille interne, le choc étant souvent assimilé à un son d'amplitude élevée et provoquant la destruction totale ou partielle des cellules ciliées. Par ailleurs, s'il est accompagné d'une fracture de l'os temporal, un traumatisme crânien peut provoquer une atteinte physique directe de la cochlée. Enfin, certains barotraumatismes peuvent aboutir à une fistule périlymphatique, qui se traduit par l'écoulement du liquide contenu dans le compartiment périlymphatique de l'oreille interne vers les cavités de l'oreille moyenne.

Dégénérescence. Plusieurs formes de dégénérescence des organes de l'oreille interne peuvent déboucher sur une surdité de perception. Ainsi l'otospongiose cochléaire (otospongiose de second type) se traduit par une calcification du labyrinthe osseux de l'oreille interne et fait partie des causes possibles de surdité endocochléaire. Par ailleurs, le vieillissement naturel de l'organisme peut parfois déboucher sur une arthérosclérose des vaisseaux cochléo-vestibulaires, également susceptible de provoquer une surdité endocochléaire. Enfin, certaines affections de type auto-immune sont susceptibles de déclencher des réactions immunitaires aboutissant à la destruction des antigènes de la cochlée et donc à une surdité neuro-sensorielle.

Surdités d'origine infectieuse, virale ou inflammatoire. La surdité de perception peut parfois remonter à la période périnatale, comme dans certains cas d'infections pendant la grossesse comme la rubéole, la syphilis ou le cytomégalovirus, ou encore en cas de prématurité, d'anoxie ou d'ictère néo-natal. Depuis la petite enfance jusqu'à l'âge adulte, l'otite, la méningite, la rougeole, les oreillons, la grippe, le zona et la polyomyélite font partie des affections dont certaines complications peuvent mener à une surdité de perception.

Maladie de Ménière. La maladie (ou syndrome) de Ménière est une maladie chronique de l'oreille interne dont la cause exacte demeure inconnue. Elle prend la forme d'épisodes imprévisibles et récurrents de vertiges, qui s'accompagnent d'acouphènes et d'une baisse d'audition. Généralement, une seule oreille est atteinte. Les symptômes sont causés par une augmentation de pression de l'endolymphe dans l'oreille interne susceptible de provoquer une destruction des cellules ciliées au fur et à mesure des crises, et donc de déboucher sur une surdité endocochléaire.

Surdités ototoxiques. Les antibiotiques de type aminosides ou glycopeptides, certains antipaludéens ou diurétiques ainsi que d'autres médicaments peuvent s'avérer toxiques pour

les structures neurosensorielles de l'oreille interne (ototoxicité[†]) et provoquer une surdité de perception. Des composés chimiques comme le plomb, les sels de mercure, le benzol, l'aniline et les composés tétrachlorés sont également ototoxiques.

Surdités rétrocochléaires

Les surdités rétrocochléaires sont des surdités de perception caractérisées par une atteinte des voies nerveuses situées après la cochlée. Lorsque cette atteinte porte sur le nerf auditif, elle peut être d'ordre toxique (alcool, tabac, certains antibiotiques ou composés chimiques ototoxiques), infectieux (méningo-névrites des oreillons, du zona, de la méningite ou de la syphilis) ou auto-immune (sclérose en plaques). Les surdités rétrocochélaires peuvent également être causées par une tumeur de l'angle ponto-cérébelleux (neurinome de l'acoustique), qui se traduit par une surdité unilatérale de perception sans recrutement, ou encore par une tumeur du tronc cérébral.

6.4 Effets de la surdité sur la perception de la parole

Sur la base des données de la littérature, nous décrivons ici les conséquences fonctionnelles d'une déprivation auditive sur la compréhension de la parole auditive, visuelle et audiovisuelle. On trouvera une description de l'évolution de la compréhension de la parole après une implantation cochléaire dans « *Effets de l'implantation sur la perception de la parole* », sect. 9.1, p. 192. Pour approfondir le sujet, nous abordons également le détail des conséquences neurofonctionnelles de la surdité sur le traitement de bas niveau dans les modalités visuelle et somesthésique dans « *Neuroplasticité intermodale après déprivation auditive* », sect. 8.3, p. 171.

6.4.1 Perception auditive de la parole lors de la surdité

En cas de déficience auditive, on peut observer une perte de la sélectivité fréquentielle (Carney et Nelson, 1983; Festen et Plomp, 1983) qui se traduit par un élargissement des largeurs de bande des filtres auditifs (Tyler *et al.*, 1984), ainsi qu'une modification de la forme de ces filtres auditifs (Glasberg et Moore, 1986; Dubno et Dirks, 1989) (voir « *Filtres auditifs et masquage psychoacoustique* », § 1.2.3, p. 27).

Pour référence, concernant le sujet normo-entendant, on pourra se référer à « *Substrats neuronaux de la parole auditive* », sect. 4.2, p. 81 pour une description des mécanismes neuronaux impliqués dans la perception et la compréhension de la parole auditive.

6.4.2 Perception visuelle de la parole lors de la surdité

Les études de la perception visuelle de la parole en cas de surdité tendent à indiquer une corrélation positive entre la durée de surdité et le niveau de lecture labiale. Ces résultats ont été reproduits aussi bien dans le cas des adultes sourds (Tillberg *et al.*, 1996) que des enfants sourds (Bergeson *et al.*, 2005).

Pour référence, concernant le sujet normo-entendant, on pourra se référer à « *Substrats neuronaux de la parole visuelle* », sect. 4.3, p. 82 pour une description des mécanismes neuronaux impliqués dans la perception et la compréhension de la parole visuelle.

6.4.3 Perception audiovisuelle de la parole lors de la surdité

De façon générale, pour les personnes normo-entendantes comme pour les personnes souffrant d'une déficience auditive, une stimulation par de la parole audiovisuelle permet d'obtenir une amélioration de la compréhension en comparaison avec une stimulation auditive ou visuelle seule.

Dans le cas des enfants déficients auditifs, plusieurs études démontrent une amélioration de la compréhension de la parole en condition audiovisuelle par comparaison avec les conditions unimodales (Erber, 1971, 1972, 1975, 1979; Arnold et Kopsel, 1996). Dans le cas particulier des enfants atteints d'une déficience auditive profonde, on n'observe pas de différence significative entre les taux d'identifications des consonnes dans les conditions audiovisuelle et visuelle, tandis qu'une différence apparaît en faveur de la stimulation audiovisuelle lors de l'utilisation de mots complets (Erber, 1972).

Pour référence, concernant le sujet normo-entendant, on pourra se réferer à *« Perception audiovisuelle de la parole »*, ch. 3, p. 59 et à *« Réseau cérébral audio-visuo-moteur de la parole »*, sect. 4.4, p. 84 pour une description des mécanismes fonctionnels et neuronaux impliqués dans la perception et la compréhension de la parole audiovisuelle.

Parole incongruente

La présentation de mots audiovisuels incongruents montre que les enfants obtenant de moins bons scores auditifs (enfants sourds) sont plus dépendants des indices visuels en cas d'incongruence audiovisuelle, tandis que les enfants obtenant de meilleurs scores auditifs (enfants normo-entendants) sont plus dépendants des indices auditifs en cas d'incongruence (Seewald *et al.*, 1985). D'après Bergeson *et al.* (2005), cette observation expérimentale suggère que la modalité sensorielle dominante pour la perception de la parole chez les enfants normo-entendants est l'audition, tandis que c'est la vision qui domine dans le cas des enfants déficients auditifs.

6.5 Prothèses auditives[2]

6.5.1 Prothèses ossiculaires

Les prothèses ossiculaires jouent un rôle de substitution pour les osselets. Elles sont indiquées dans certains cas d'atteinte de l'oreille moyenne. Ces atteintes sont pour la plupart d'origine inflammatoire (otite chronique éventuellement accompagnée d'une lyse ossiculaire), dégénérative (otospongiose), malformative ou traumatique (luxation des articulations ossiculaires, éventuellement accompagnée d'une fracture ossiculaire).

Les atteintes inflammatoires de l'oreille moyenne sont le plus souvent soignées par un traitement médicamenteux mais nécessitent parfois une intervention chirurgicale. Lorsque la chaîne des osselets est détériorée, il est possible de rétablir la continuité de cette chaîne soit en remodelant les osselets (ossiculoplastie), soit en remplaçant certains d'entre eux par des prothèses ossiculaires. En cas d'otospongiose, qui provoque une ankylose de l'étrier, on remplace le corps de l'étrier par un piston millimétrique en téflon.

6.5.2 Aides auditives externes

En cas de surdité neurosensorielle, il est souvent possible de proposer à la personne malentendante une prothèse auditive externe, qui se loge dans le conduit auditif et permet une

[2] d'après Bouccara *et al.* (2005)

amplification sélective des sons extérieurs. De telles prothèses sont proposées au patient après une évaluation multidisciplinaire associant le médecin traitant et l'oto-rhino-laryngologiste avec un audioprothésiste, un orthophoniste et un psychologue. Cette évaluation a pour but de quantifier précisément la déficience auditive et ses conséquences sur le plan quotidien.

De nos jours, la plupart des prothèses auditives externes sont numériques. Cette technologie présente plusieurs avantages importants :

- multiprogrammation : en fonction des différents milieux sonores rencontrés dans la vie quotidienne, il est possible d'utiliser plusieurs jeux de réglages, chacun adapté à un environnement sonore particulier ;
- identification de la voix : le processeur amplifie de façon sélective les canaux fréquentiels spécifiques de la voix humaine, afin d'améliorer le rapport signal/bruit. Il est également possible de réduire l'amplification en l'absence de signaux vocaux afin d'atténuer les bruits non pertinents ;
- compression multicanaux : en cas de bruits intenses, le processeur numérique réduit l'amplification dans les canaux fréquentiels concernés, sans modifier l'amplification des autres canaux fréquentiels.

6.5.3 Implants d'oreille moyenne

Les implants d'oreille moyenne permettent de stimuler directement la chaîne ossiculaire par des vibrations mécaniques. Ils sont indiqués dans le cas où les prothèses externes conventionnelles ne peuvent pas être utilisées (eczéma, otite externe, sténose) ou lorsque ces dernières ne parviennent pas à restaurer convenablement les capacités auditives.

Le dispositif est constitué d'un microphone destiné à capter les sons de l'environnement, un processeur permettant un traitement fin des signaux sonores, et un transducteur électromagnétique qui convertit les signaux numériques en vibrations mécaniques. La stimulation mécanique est assurée par l'intermédiaire d'un vibrant fixé chirurgicalement sur l'enclume.

6.5.4 Neuroprothèses auditives

Les neuroprothèses auditives regroupent les implants cochléaires et les implants du tronc cérébral. Ces neuroprothèses sont indiquées en dernier recours, dans le cas où aucun autre type de prothèse ne parvient à restaurer un niveau de compréhension de la parole acceptable, en particulier dans le cas des surdités bilatérales sévères à profondes. En raison de leur prix élevé, les neuroprothèses sont usuellement implantées d'un seul côté.

Une description complète de l'implant cochléaire est proposée dans *« L'implant cochléaire »*, ch. 7, p. 145. Dans le cas où l'implantation cochléaire est impossible, par exemple lorsque la cochlée ou le nerf auditif sont profondément détériorés, il est possible d'utiliser un implant du tronc cérébral, la stimulation électrique étant dans ce cas permise par l'implantation d'un porte-électrodes dans les noyaux cochléaires du tronc cérébral. En dépit de cette différence, le principe de fonctionnement de l'implant du tronc cérébral est tout à fait similaire à celui de l'implant cochléaire.

6.5.5 Prothèses à ancrage osseux

Les prothèses à ancrage osseux permettent de stimuler mécaniquement l'oreille interne en mettant en vibration les os de la voûte crânienne. Elles sont indiquées en cas d'atteinte de l'oreille moyenne, l'oreille interne restant préservée fonctionnellement, ainsi qu'en cas de surdité totale

unilatérale, les vibrations en provenance du côté atteint permettant de restaurer partiellement l'audition bilatérale.

La stimulation des os de la voûte crânienne se fait grâce à un pilier en titane inséré chirurgicalement dans l'os. Une prothèse auditive externe de type conventionnel est alors adaptée à ce pilier, et transmet sous forme de vibrations mécaniques les informations sonores extérieures.

6.6 Communication avec les personnes sourdes

6.6.1 La langue des signes

La langue des signes est une langue visuelle. C'est le moyen de communication qu'utilisent les personnes sourdes pour dialoguer entre elles. Elle est une langue à part entière au même titre que les langues parlées telles que le français ou l'anglais. Les professionnels en contact avec des déficients auditifs la pratiquent, comme par exemple les enseignants, les orthophonistes, les travailleurs sociaux ou les interprètes spécialisés.

Un signe est caractérisé par cinq paramètres : positions des doigts et de la main, mouvements, emplacement et expressions du visage. Les signes sont structurés entre eux par une syntaxe (lieu, protagonistes, action) et une grammaire. Il n'existe pas de langage des signes universel, car chaque pays a son propre langage des signes, mais les sourds de pays différents communiquent facilement entre eux après un petit temps d'adaptation. On distingue notamment la LSF (langue des signes française), le BSL (british sign language), l'ASL (american sign language), le QSL (Quebec sign language), etc. En France, la LSF est actuellement pratiquée par plusieurs centaines de milliers de personnes sourdes.

6.6.2 Méthodes oralistes de communication

La lecture labiale

La lecture labiale (*lipreading*), ou lecture oro-faciale (*speechreading*), consiste à observer les mouvements des lèvres, de la langue et du visage d'un interlocuteur que l'on est dans l'impossibilité d'entendre afin de comprendre ce qu'il dit. L'appellation de *lecture labiale* traditionnellement utilisée est en fait inappropriée, puisque des éléments anatomiques comme la langue, les muscles du visage ou les mouvements de la tête sont également susceptibles de fournir des indices sur le contenu phonétique et sémantique du discours. Il serait donc plus rigoureux de parler de lecture oro-faciale, ou *speechreading* en anglais.

Toutefois l'utilisation des seules informations visuelles ne préserve pas toute l'information phonétique présente dans la parole, en raison des homonymies labiales : par exemple les familles de phonèmes [/p/ /b/ /m/], [/t/ /d/ /n/], [/k/ /g/], [/s/ /z/], [/ʃ/ /ʒ/] ou [/f/ /v/] sont pratiquement impossibles à différencier par la seule lecture labiale. D'autre part plusieurs phonèmes ont une articulation postérieure, ce qui les rend difficilement discriminables (comme par exemple /t/ et /k/).

En raison de ces ambigüités visuelles, l'identification correcte des mots prononcés dépend dans une proportion non négligeable du contexte syntaxique, grammatical et sémantique dans lequel ces mots viennent s'insérer. Par conséquent un apprentissage convenable de la lecture labiale doit reposer sur un entraînement de long terme, réalisé dans des conditions réalistes de communication en face-à-face.

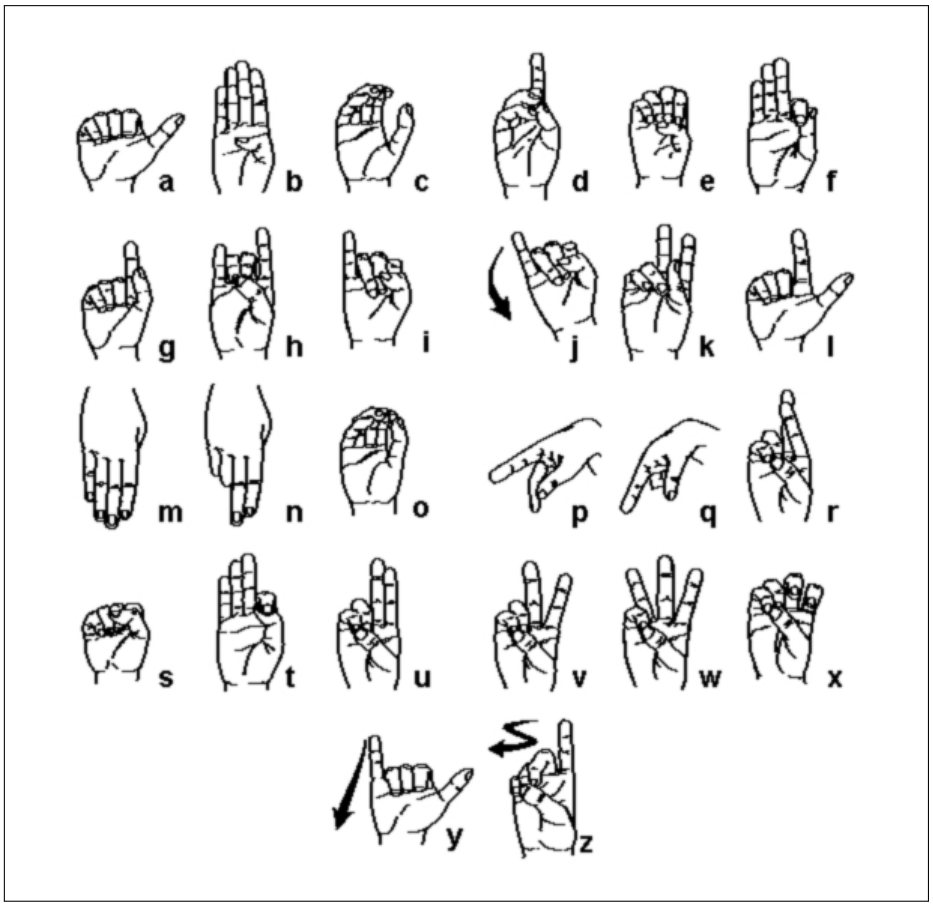


FIG. 6.2 – Codage des lettres de l'alphabet français en LSF

Le langage parlé complété

Mis au point aux Etats-Unis par le docteur R. Orin Cornett en 1967 (ALPC, 2004), le langage parlé complété — *cued speech* en anglais — s'est d'abord développé dans les pays anglophones, et est aujourd'hui adapté à plus de 40 langues dans le monde. Introduit en France en 1977, le LPC connait depuis un développement spectaculaire auprès des familles, des professionnels et de nombreux centres spécialisés pour déficients auditifs. Le LPC est une aide à la lecture labiale, dont il supprime les ambiguïtés en permettant de différencier les phonèmes visuellement identiques. Il consiste à associer à chaque phonème prononcé un geste de complément effectué par la main près du visage, qui complète syllabe par syllabe tout ce que le locuteur prononce. Chaque syllabe se code en mettant la main à la position correspondant à la voyelle, les doigts réalisant la clé de la consonne. L'acquisition des structures de la langue grâce au LPC permet aux enfants sourds d'apprendre à lire à un âge normal et en suivant les mêmes étapes qu'un enfant entendant.

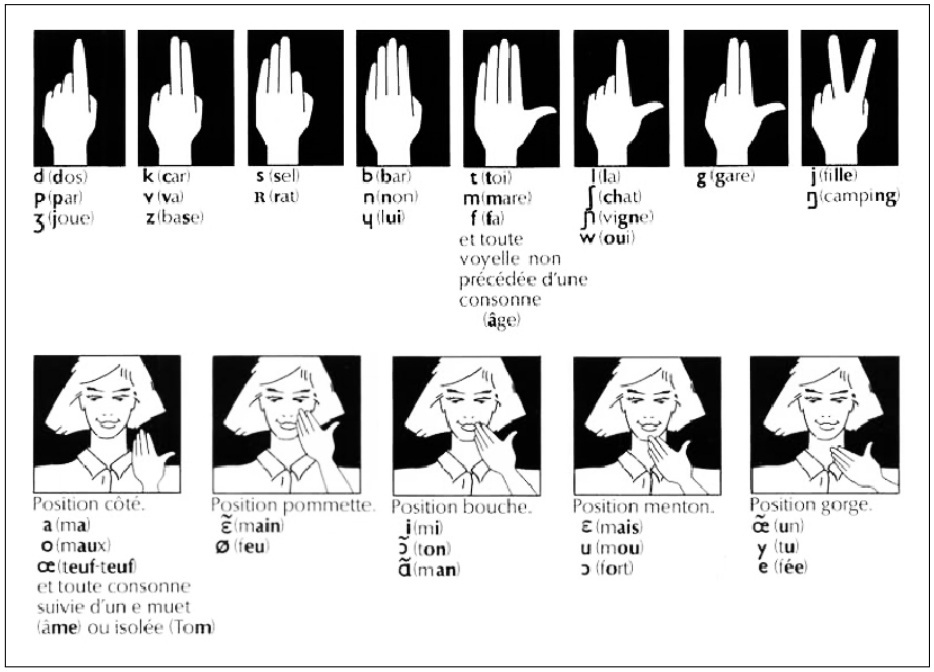


FIG. 6.3 – Code LPC pour les consonnes et voyelles françaises

Bibliographie

ALPC (2004). Le code LPC. Association pour la promotion et le développement du langage parlé complété, site internet. http://www.alpc.asso.fr/code01-c.htm.

Arnold, P. et Kopsel, A. (1996). Lipreading, reading and memory of hearing and hearing-impaired children. *Scand. Audiol.*, 25(1):13–20.

Bergeson, T. R., Pisoni, D. B. et Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear Hear.*, 26(2):149–164.

BIAP (1997). Recommandation BIAP 02-1 bis : classification audiométrique des déficiences auditives. Bureau international d'audiophonologie, site internet. http://www.biap.org/recom02-1.htm.

Bouccara, D., Avan, P., Mosnier, I., Bozorg Grayeli, A., Ferrary, E. et Sterkers, O. (2005). Réhabilitation auditive. *Med. Sci.*, 21(2):190–197.

Carney, A. E. et Nelson, D. A. (1983). An analysis of psychophysical tuning curves in normal and pathological ears. *J. Acoust. Soc. Am.*, 73(1):268–278.

Dubno, J. R. et Dirks, D. D. (1989). Auditory filter characteristics and consonant recognition for hearing-impaired listeners. *J. Acoust. Soc. Am.*, 85(4):1666–1675.

Erber, N. P. (1971). Auditory and audiovisual reception of words in low-frequency noise by children with normal hearing and by children with impaired hearing. *J. Speech Hear. Res.*, 14(3):496–512.

Erber, N. P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *J. Speech Hear. Res.*, 15(2):413–422.

Erber, N. P. (1975). Auditory-visual perception of speech. *J. Speech Hear. Disord.*, 40(4):481–492.

Erber, N. P. (1979). Speech perception by profoundly hearing-impaired children. *J. Speech Hear. Disord.*, 44(3):255–270.

Festen, J. M. et Plomp, R. (1983). Relations between auditory functions in impaired hearing. *J. Acoust. Soc. Am.*, 73(2):652–662.

Glasberg, B. R. et Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J. Acoust. Soc. Am.*, 79(4):1020–1033.

Granier-Deferre, C., Lecanuet, J. P., Cohen, H. et Busnel, M. C. (1985). Feasibility of prenatal hearing test. *Acta Otolaryngol. Suppl.*, 421:93–101.

O'Donoghue, G. M. (1996). Cochlear implants in children : principles, practice and predictions. *J. R. Soc. Med.*, 89(6):345.

O'Donoghue, G. M., Nikolopoulos, T., Archbold, S. M. et Tait, M. (1998). Congenitally deaf children following cochlear implantation. *Acta Otorhinolaryngol. Belg.*, 52(2):111–114.

Seewald, R. C., Ross, M., Giolas, T. G. et Yonovitz, A. (1985). Primary modality for speech perception in children with normal and impaired hearing. *J. Speech Hear. Res.*, 28(1):36–46.

Tillberg, I., Ronnberg, J., Svard, I. et Ahlner, B. (1996). Audio-visual speechreading in a group of hearing aid users. The effects of onset age, handicap age, and degree of hearing loss. *Scand. Audiol.*, 25(4):267–272.

Tyler, R. S., Hall, J. W., Glasberg, B. R., Moore, B. C. et Patterson, R. D. (1984). Auditory filter asymmetry in the hearing impaired. *J. Acoust. Soc. Am.*, 76(5):1363–1368.

Chapitre 7

L'implant cochléaire

Sommaire

7.1 Principe 145
7.2 Indications et critères d'implantation 147
7.2.1 Indications 147
7.2.2 Critères audiométriques 147
7.2.3 Critères oto-rhino-laryngologiques 147
7.2.4 Critères psychologiques et environnementaux 148
7.3 Implantation cochléaire 148
7.4 Rééducation fonctionnelle 149
7.5 Aspects techniques 149
7.5.1 Traitement du signal 149
7.5.2 Electrodes 151
Bibliographie 153

Introduction

Une caractéristique exceptionnelle des implants cochléaires est qu'ils fournissent un substitut complet du système auditif moyen et externe en établissant directement une interface entre les centres nerveux auditifs d'une personne et son environnement sonore extérieur. Issus de la rencontre multidisciplinaire entre la recherche biomédicale, les neurosciences et l'ingénierie électronique appliquée au traitement du signal, les implants cochléaires ont été développés à partir de la fin du XX^{e} siècle, leurs performances techniques et leur champ d'application subissant une progression constante. A l'heure actuelle, on dénombre de par le monde plus de 100 000 porteurs d'implants cochléaires, dont environ 30 000 enfants.

7.1 Principe

Un implant cochléaire est un appareil destiné à rétablir partiellement l'audition chez les personnes atteintes de surdité sévère ou profonde, grâce à une stimulation électrique adéquate des fibres nerveuses auditives résiduelles de la cochlée. L'implant est indiqué dans les cas où l'organe

de Corti ne s'est pas développé ou a été endommagé, rendant impossible la restitution d'une audition satisfaisante par des prothèses auditives externes. Pour des précisions supplémentaires sur la surdité et le fonctionnement du système auditif, on pourra se référer à *« La surdité »*, ch. 6, p. 127, et à *« Système auditif périphérique »*, sect. 1.1, p. 18.

Un implant cochléaire est constitué de deux parties fonctionnelles interdépendantes. La partie interne, destinée à être implantée chirurgicalement, est composée d'une antenne de réception, d'un récepteur-stimulateur et d'un porte-électrode. La partie externe comprend un microphone permettant de capter l'information acoustique, un processeur vocal, une antenne de transmission et le système d'alimentation énergétique.

Lorsque le microphone de la partie externe capte un signal acoustique, il le convertit en un signal électrique qui est alors transmis au processeur vocal. Celui-ci filtre électroniquement le signal envoyé par le microphone et le convertit en une série de signaux électriques qui sont envoyés à travers la peau par l'antenne externe sous forme d'un signal radiomodulé en fréquence. L'antenne de réception placée chirurgicalement sous la peau décode ce signal et le transmet aux électrodes correspondantes, ce qui permet de stimuler électriquement les fibres du nerf auditif reliées aux cellules ciliées de la cochlée. Les porte-électrodes modernes comportent jusqu'à 24 électrodes stimulatrices destinées à être placées dans la rampe tympanique de la cochlée. Ils sont soit droits, soit enroulés à mémoire de forme, ce qui permet d'assurer un meilleur contact entre les électrodes et le tissu neural.

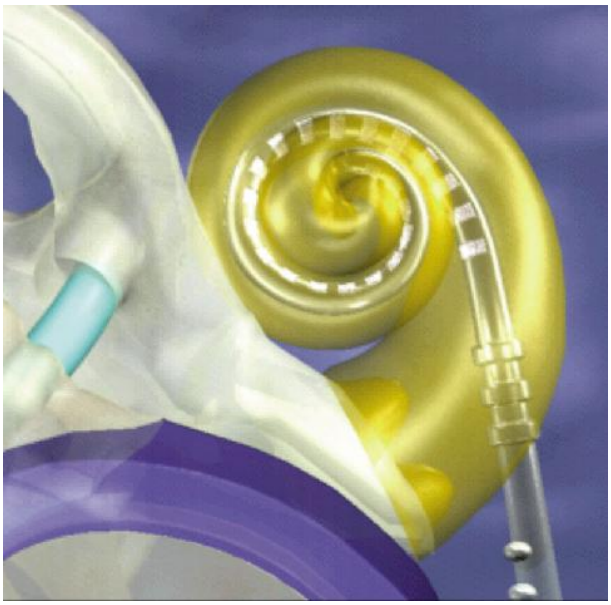


FIG. 7.1 – Infographie montrant le porte-électrodes de l'implant cochléaire inséré dans la rampe tympanique de la cochlée.

Il existe actuellement quatre fabricants d'implants cochléaires : MXM en France, commercialisant l'implant Digisonic ; Advanced Bionics Corporation en Amérique du Nord, fabriquant l'implant Clarion ; Cochlear Corporation en Australie (avec une participation nord-américaine), commercialisant l'implant Nucleus ; MedEl Corporation en Autriche, commercialisant l'implant Combi 40.

7.2 Indications et critères d'implantation

7.2.1 Indications

Les implants cochléaires ont à l'origine été mis au point afin de traiter les pertes auditives liées à un organe de Corti dysfonctionnel. Toutefois les indications d'implantation cochléaire ont été étendues progressivement durant les quinze dernières années, et elles incluent à présent les personnes possédant une audition résiduelle mais ne tirant aucun bénéfice des prothèses acoustiques conventionnelles (von Ilberg *et al.*, 1999; Gantz et Turner, 2003; Kiefer *et al.*, 2004), les personnes souffrant d'un acouphène chronique (Thedinger *et al.*, 1985; McKerrow *et al.*, 1991; Dauman *et al.*, 1993; Ito et Sakakihara, 1994; Tyler, 1995; Ruckenstein *et al.*, 2001), ainsi que celles souffrant d'une neuropathie auditive (Miyamoto *et al.*, 1999; Trautwein *et al.*, 2000; Shallop *et al.*, 2001; Sininger et Trautwein, 2002; Peterson *et al.*, 2003).

Si le patient est atteint d'une surdité prélinguale, il est fortement recommandé d'effectuer l'implantation le plus tôt possible avant la période d'acquisition du langage. En effet les patients sourds prélinguaux ayant subi une implantation cochléaire pendant leur adolescence ou à l'âge adulte rencontrent d'importantes difficultés pour accéder à une compréhension de la parole (Tyler et Lowder, 1992). On pourra en particulier se référer à *« Période sensible pour l'implantation cochléaire »*, § 9.3.3, p. 199 pour plus de détails sur les observations comportementales, électrophysiologiques et neurofonctionnelles justifiant une implantation précoce en cas de surdité prélinguale.

7.2.2 Critères audiométriques

Les critères audiométriques de sélection des patients pour une implantation cochléaire ont évolué en parallèle avec les progrès techniques et chirurgicaux dans le domaine ainsi que les observations cliniques des progrès perceptifs chez les patients implantés. A la naissance de l'implant cochléaire, dans les années 1980, le critère audiométrique initial était une surdité totale bilatérale, caractérisée par une perte auditive supérieure à 110 dB. Ce critère a été élargi dans les années 1990 aux déficiences auditives sévères, caractérisées par une perte auditive supérieure à 70 dB (pour plus de détails techniques concernant l'audiométrie, on pourra consulter *« ?? »*, § ??, p. ??). Les critères employés actuellement ne sont plus évalués en termes de perte auditive mais plutôt en termes de compréhension de la parole, et on considère comme candidat potentiel pour une implantation cochléaire un patient dont le score perceptif de compréhension des phrases en liste ouverte ne dépasse pas 50% avec des prothèses acoustiques externes conventionnelles.

7.2.3 Critères oto-rhino-laryngologiques

Si l'audiogramme tonal ne met en évidence aucun reste auditif, on peut recourir au test de la fenêtre ronde, qui permet d'évaluer l'état du nerf auditif en donnant une approximation du pourcentage de fibres nerveuses encore en état de fonctionnnement (Shipp et Nedzelski, 1994). Ce test est réalisé sous anesthésie locale et sans hospitalisation chez l'adulte et sous anesthésie générale chez l'enfant. Il consiste à stimuler électriquement la membrane de la fenêtre ronde, ce qui provoque la diffusion du stimulus électrique dans les liquides cochléaires et permet d'atteindre les terminaisons nerveuses de la cochlée.

La sensation sonore éventuellement obtenue permet d'attester la positivité du test. Les niveaux liminaires qui permettent d'obtenir une réponse sont soigneusement évalués. L'enregistrement à différentes fréquences de ces seuils permet de déterminer objectivement l'état du nerf auditif. En cas de test positif, la réhabilitation par implantation cochléaire est envisageable.

Dans tous cas, il faut s'assurer de la bonne perméabilité du tube cochléaire. Le diagnostic est la plupart du temps réalisé par imagerie par résonance magnétique, ce qui permet de visualiser les liquides du labyrinthe. Il faut également tenir compte des difficultés opératoires que représenterait une pathologie locale associée à la surdité, comme une cavité d'évidement ou une otite chronique évolutive.

De façon générale, le choix du côté à implanter dépend directement du degré d'ossification cochléaire, de la durée de surdité avant implantation, des résultats de stimulation électrique de la fenêtre ronde, et indirectement de la dominance cérébrale (Deguine *et al.*, 1995).

7.2.4 Critères psychologiques et environnementaux

Ces critères sont particulièrement importants et peuvent amener l'équipe médicale à refuser l'implantation cochléaire. Les motivations du patient, sa profession et ses possibilités de réinsertion socio-professionnelle, son niveau intellectuel, ses qualités relationnelles et comportementales sont essentielles. Par exemple, la revendication de retrouver une audition injustement perdue, le souci d'obtenir à nouveau une audition normale ou chez l'enfant, d'acquérir à brève échéance le langage parlé et surtout, l'attente d'un miracle sont des contre-indications à l'implantation cochléaire. C'est lors des entretiens préalables et du bilan préopératoire que sont évalués l'aptitude du patient à tirer profit de l'implant et les niveaux d'un certain nombre d'éléments comme la voix, l'articulation, le langage et la lecture labiale. Ceux-ci serviront de référence après l'intervention pour évaluer les résultats obtenus.

7.3 Implantation cochléaire[1]

L'implantation cochléaire nécessite une anesthésie générale pour mettre en place la partie interne de l'implant. L'intervention chirurgicale peut durer de 1h30 à 4h selon les circonstances. Le récepteur-stimulateur et l'antenne réceptrice de la partie interne sont placés sous la peau par une incision derrière le pavillon de l'oreille, une logette pouvant être creusée dans l'os par fraisage pour assurer leur placement convenable.

On réalise ensuite une mastoïdectomie puis une tympanotomie postérieure pour accéder à la caisse du tympan. La tympanotomie est réalisée entre la troisième portion interpétreuse du nerf facial et l'insertion de la membrane tympanique. Il est préférable d'effectuer un monitorage peropératoire du nerf facial afin de minimiser toute agressivité vis-à-vis de ce dernier. Une fois la caisse du tympan atteinte, on réalise une cochléostomie, qui permettra l'insertion du porte-électrodes dans la scala tympani. On obture ensuite de manière étanche la cochlée au moyen du fragment libre de facia temporalis et de muscle temporal de façon à éviter toute fistule labyrinthique, qui pourrait exposer le patient à un risque accru d'acouphènes ou de vertiges post-opératoires et également de méningite bactérienne en cas de survenue d'une otite moyenne aigüe.

En cas de cochlée ossifiée, l'intervention est beaucoup plus complexe et ses résultats plus aléatoires. Ces difficultés sont directement liées à l'importance de l'obstruction, et à la localisation de celle-ci sur la longueur du clavier cochléaire. Le cas le plus grave est l'obstruction totale, car on ne peut introduire le porte-électrode. On est obligé d'utiliser un système à électrodes séparées placées une à une dans des puits forés dans l'os cochléaire.

Les complications de l'implantation cochléaire sont actuellement bien documentées. Les problèmes de cicatrisation du lambeau apparaissent dans environ 2% des cas, ce type de complication se raréfiant avec le développement de techniques chirurgicales faisant appel à de simples

[1] d'après Truy et Lina (2003)

incisions. Une paralysie faciale peut survenir dans 0,65 à 0,93% des cas lorsque le nerf facial a été touché. Le mauvais placement de l'électrode par des chirurgiens manquant d'expérience se retrouve dans 0,80% des cas. On peut rencontrer un tassement et une mise en court-circuit de quelques électrodes intracochléaires, ce qui est en général sans conséquence dans la mesure où un certain nombre d'électrodes restent fonctionnelles (0,5% des cas). Les autres complications sont une oreille geyser, c'est-à-dire une expulsion de liquide céphalorachidien lors de la cochléostomie, en général dans un contexte d'oreille malformative (0,05% à 0,5% des cas) et une extrusion du porte-électrodes (0,13%).

7.4 Rééducation fonctionnelle

L'implantation cochléaire nécessite une prise en charge post-opératoire pour pouvoir tirer entièrement profit de l'appareillage. Les besoins de rééducation de chaque patient dépendent de leur expérience auditive préopératoire. Pour le patient sourd prélingual, des entraînements pour l'audition et la parole sont impératifs pour améliorer leurs capacités à communiquer correctement. Pour le patient sourd postlingual, un entraînement plus complexe est souvent requis. L'implantation cochléaire chez les enfants est réussie lorsqu'elle est suivie par un entraînement intensif par une équipe de réhabilitation multidisciplinaire, dont le but est de permettre à l'enfant malentendant d'apprendre passivement à partir de son environnement. La réhabilitation prend en compte la compréhension du langage parlé et l'expression orale.

La prise en charge post-opératoire comporte un suivi médical, un entraînement au décodage de la parole, un apprentissage de l'écoute musicale, des séances de réglage de l'appareil, des tests successifs d'évaluation des performances des patients ainsi qu'un important soutien psychologique. L'entraînement auditif à proprement parler comportera une adaptation aux manipulations de l'émetteur, une initation à l'environnement sonore, une reconnaissance des rythmes, des intensités, des fréquences, de la mélodie, la découverte et l'analyse des nouvelles perceptions musicales, un entraînement à la discrimination ou à la reconnaissance de la parole ainsi qu'un entraînement ou une reprise de la lecture labiale.

7.5 Aspects techniques

7.5.1 Traitement du signal

Le rôle principal du processeur externe de l'implant est de transformer les signaux sonores en signaux électriques cohérents. Les implants actuels multi-électrodes tirent tous partie de l'organisation tonotopique naturelle de la cochlée en découpant le signal acoustique original suivant différentes bandes de fréquences, puis en transmettant les hautes fréquences vers la partie basale de la cochlée et les basses fréquences vers la partie apicale de la cochlée. Toutefois les stratégies de traitement intervenant dans l'extraction, le codage et la stimulation électrique dans ces différentes bandes de fréquence diffèrent suivant les constructeurs.

Ainsi, en utilisant le même implant avec différentes stratégies de codage électrique, il est possible d'observer des différences majeures dans la compréhension de la parole. Les premières recherches ont montré qu'on observait lors de la stimulation électrique du nerf auditif un phénomène d'« entonnoir » électro-neural, limitant la quantité d'information acoustique pouvant être présentée au nerf auditif. Pour outrepasser ce phénomène, il est donc préférable d'adopter une stratégie de codage qui ne transmet que les informations de première importance pour la compréhension de la parole. On distingue ainsi les stratégies utilisant une analyse temporelle du signal acoustique, ce qui permet de transmettre au mieux l'ensemble des sons ambiants sans privilégier

spécifiquement la parole, les stratégies spectrales, qui extraient spécifiquement les informations acoustiques significatives pour la parole humaine, et les stratégies hybrides, qui utilisent une approche à la fois spectrale et temporelle.

Stratégies temporelles

Une première possibilité est de transmettre directement le signal acoustique sous forme analogique pour chaque bande de fréquence (Eddington *et al.*, 1978). Ainsi, dans le tout-premier modèle d'implant cochléaire mono-électrode (3M/House), le signal acoustique original était traité sous forme analogique, subissait une compression d'amplitude pour respecter la dynamique électrique du nerf auditif puis était transmis directement à la cochlée après une modulation d'amplitude à 16000 Hz (Zeng, 2004). Dans le cadre des implants actuels multi-électrodes, on distingue les stratégies CA (*compressed analog*) et SAS (*simultaneous analog stimulation*). Dans le codage analogique utilisé pour l'implant Ineraid (stratégie CA), l'électrode de retour est située à l'extérieur de la cochlée, dans le muscle temporal (mode monopolaire). Au contraire, dans le codage analogique utilisé pour l'implant Clarion (stratégie SAS), l'électrode de retour est systématiquement adjacente à l'électrode de stimulation (mode bipolaire).

Une seconde possibilité est d'effectuer un transcodage numérique afin de restituer fidèlement l'enveloppe temporelle dans chaque bande de fréquence, par l'intermédiaire d'un traitement de type « vocoder » (Dudley, 1939a,b). Dans ce cadre, la stratégie CIS (*continuous interleaved sampling*) stimule successivement chaque électrode dans un ordre spécifique tout en préservant les informations relatives à l'enveloppe temporelle, ce qui permet de réduire considérablement les interactions électriques entre les différents canaux (Wilson *et al.*, 1991). Cette stratégie effectue, dans chaque bande fréquentielle, une extraction d'enveloppe par rectification d'onde et filtrage passe-bas, la fréquence de coupure correspondante se situant typiquement entre 160 et 320 Hz. L'enveloppe subit ensuite une compression de type logarithmique afin de respecter la faible dynamique électrique du nerf auditif. Pour finir, l'enveloppe temporelle de chaque bande fréquentielle subit une modulation d'amplitude par une porteuse constituée d'impulsions biphasiques à haute fréquence (typiquement entre 800 et 1600 Hz). Les pulsations de chaque bande fréquentielle sont alors émises de façon séquentielle, s'entrecroisant sans aucune coïncidence temporelle, ce qui permet d'éviter des interactions entre électrodes dues à un chevauchement des champs électriques lié à une stimulation simultanée. Une implémentation typique de la stratégie CIS implique un nombre de filtres passe-bande identique au nombre d'électrodes, variant de 6 dans les premières implémentations pour l'implant Ineraid à 8 dans l'implant Clarion, 12 dans l'implant Medel et 22 ou 24 dans l'implant Nucleus.

Afin de permettre un compromis entre haute fréquence de stimulation et minimisation des interactions entre électrodes, les pulsations correspondant à des électrodes suffisamment distantes les unes des autres peuvent également être émises simultanément. Ce compromis a été mis en application dans les stratégies de codages PPS (*paired pulsatile stimulation*) et MPS (*multiple pulsatile stimulation*), développées dans le cas de l'implant Clarion, et qui utilisent respectivement des stimulations simultanées de deux ou plusieurs électrodes (Loizou *et al.*, 2003).

Stratégies spectrales

Une autre approche consiste à se placer du point de vue de la perception et la production de la parole : dans cette approche, les différents pics spectraux sont associés aux formants de la parole, qui traduisent la résonance acoustique des différentes cavités du conduit vocal (voir *« Caractéristiques acoustiques des phonèmes »*, sect. 2.3, p. 52). Il s'agit donc d'extraire et de

transmettre en priorité ces formants ou ces maxima spectraux afin de concentrer le traitement du signal sur ces informations cruciales dans la compréhension de la parole.

Cette approche a été tout d'abord mise en oeuvre dans la stratégie par second formant F0-F2, implémentée dans le processeur Nucleus WSP II (Clark *et al.*, 1984). Cette stratégie consiste à extraire les fréquences du fondamental F0 et du second formant F2 (essentiel pour la compréhension de la parole), F2 déterminant la position de l'électrode de stimulation et F0 la fréquence de stimulation de cette électrode. Le premier formant F1 a ensuite été introduit, dans la stratégie par premier et second formants F0-F1-F2, implémentée dans le processeur Nucleus WSP III, avec une méthode de stimulation similaire. Enfin, une extraction des principaux pics spectraux a été mise en place dans la stratégie de codage MPEAK, développée par Nucleus (Skinner *et al.*, 1991).

Stratégies hybrides

Les stratégies de codage les plus récentes essaient d'intégrer le meilleur traitement en mettant l'accent à la fois sur l'analyse de la structure spectrale et de l'information temporelle. La désignation générique de ces stratégies, « n of m », signifie qu'un certain nombre de canaux (n) parmi un nombre total (m) sera stimulé à un instant donné. Le nombre maximum de stimulations est prédéterminé par l'audiologiste ou par le technicien chargé de régler l'implant, le nombre d'électrodes réellement stimulées à un instant donné étant susceptible de varier dans la limite de ce maximum. Dans tous les cas, le codage privilégie les bandes fréquentielles avec les plus fortes énergies spectrales. Ce type de codage a été utilisé pour la première fois dans le processeur Nucleus SPEAK (McDermott *et al.*, 1992).

De manière à améliorer la représentation spectrale et temporelle du signal acoustique, la fréquence d'analyse de la parole a été améliorée dans les stratégies les plus récentes de codage, avec des fréquences de stimulation élevées (au-delà de 2000 Hz). Les stratégies hybrides à haute fréquence sont actuellement HiResolution dans l'implant Clarion, ACE dans l'implant Nucleus, et Tempo dans l'implant MedEl.

7.5.2 Electrodes

Les électrodes des premiers implants cochléaires étaient composées de cuivre et d'or, tandis que de nos jours elles sont composées d'alliages de platine et d'iridium. Les premiers implants ne comportaient qu'une électrode (House/3M et Vienna/3M) ; les implants multi-électrodes actuels comportent quant à eux de 6 électrodes pour l'implant Ineraid à 24 électrodes pour l'implant Nucleus CI-24. L'espacement inter-électrode peut être constant (de 0.75 mm pour l'implant Nucleus à 4 mm pour l'implant Ineraid) ou se réduire progressivement de la base vers l'apex de la cochlée (implant Clarion). Les formes des électrodes diffèrent selon les constructeurs, avec une structure annulaire ou semi-annulaire dans l'implant Nucleus, sphérique dans l'implant Clarion et bisphérique dans l'implant Medel. Le positionnement des électrodes dans la cochlée varie en fonction du déroulement de l'intervention chirurgicale et des caractéristiques physiologiques de la cochlée, avec des inégalités possibles dans la profondeur d'insertion, des torsions ou des retournements possibles du porte-électrode, et dans certains cas une perforation de la membrane basilaire se traduisant par une arrivée du porte-électrodes dans la rampe vestibulaire (Skinner *et al.*, 2002).

Plus il y a d'électrodes, plus le nombre de bandes de fréquences stimulées est important. Toutefois, l'espacement minimal entre deux électrodes doit être suffisant pour qu'elles ne se perturbent pas entre elles. Si cette distance est abaissée, deux électrodes voisines vont stimuler les mêmes cellules ciliées et il y aura diaphonie entre les deux signaux. Les électrodes ont en fait

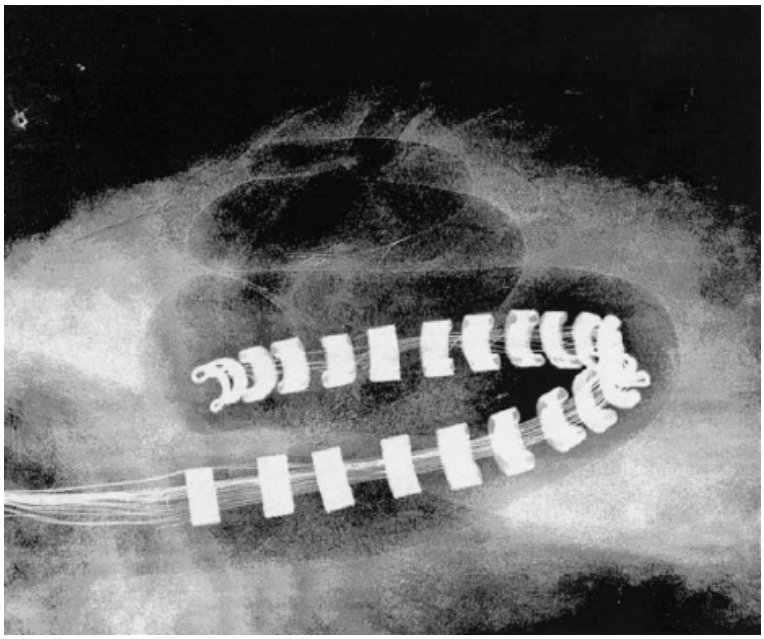


FIG. 7.2 – Implant cochléaire Nucleus CI-22 inséré dans la rampe tympanique. Adapté d'après Clark (2003).

un large rayon d'action en raison de la présence de l'endolymphe à l'intérieur de la cochlée. On distingue des modes de stimulation monopolaires, bipolaires, tripolaires et à masse commune. Dans les cas des modes monopolaires, une électrode est placée en dehors de la cochlée, en général dans le muscle temporal, et sert d'électrode de référence pour les électrodes de stimulation. Dans le cas des modes bipolaires, une électrode adjacente à l'électrode de stimulation est utilisée comme électrode de retour du courant électrique, ce qui offre des possibilités importantes de déclinaisons de ce mode, selon que l'électrode de stimulation est proche ou non de l'électrode de recueil. Dans le cas des modes tripolaires, ce sont les deux électrodes voisines de l'électrode de stimulation qui servent d'électrodes de retour. Enfin, dans le cas du mode à masse commune, toutes les électrodes non stimulantes à un instant donné servent de retour pour l'électrode stimulante. Les données issues des modélisations et des mesures physiologiques montrent une dispersion décroissante du champ électrique (et donc une précision spatiale croissante dans la stimulation électrique) selon que le mode de stimulation est respectivement monopolaire, bipolaire ou tripolaire (Jolly *et al.*, 1996; Bierer et Middlebrooks, 2002, 2004).

Bibliographie

Bierer, J. A. et Middlebrooks, J. C. (2002). Auditory cortical images of cochlear-implant stimuli : dependence on electrode configuration. *J. Neurophysiol.*, 87(1):478–492.

Bierer, J. A. et Middlebrooks, J. C. (2004). Cortical responses to cochlear implant stimulation : channel interactions. *J. Assoc. Res. Otolaryngol.*, 5(1):32–48.

Clark, G. M. (2003). Cochlear implants in children : safety as well as speech and language. *Int. J. Pediatr. Otorhinolaryngol.*, 67:1.

Clark, G. M., Tong, Y. C. et Dowell, R. C. (1984). Comparison of two cochlear implant speech-processing strategies. *Ann. Otol. Rhinol. Laryngol.*, 93(2 Pt 1):127–131.

Dauman, R., Tyler, R. S. et Aran, J. M. (1993). Intracochlear electrical tinnitus reduction. *Acta Otolaryngol.*, 113(3):291–295.

Deguine, O., Garcia de Quevedo, S., Fraysse, B., Cormary, X., Uziel, A. et Demonet, J. F. (1995). Criteria for selecting the side for cochlear implantation. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 166:403–406.

Dudley, H. (1939a). Remaking Speech. *J. Acoust. Soc. Am.*, 11(2):169–177.

Dudley, H. (1939b). The vocoder. *In Bell Labs Record*, pages 122–126.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G. et Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *Ann. Otol. Rhinol. Laryngol.*, 87(6 Pt 2):1–39.

Gantz, B. J. et Turner, C. W. (2003). Combining acoustic and electrical hearing. *Laryngoscope*, 113(10):1726–1730.

Ito, J. et Sakakihara, J. (1994). Tinnitus suppression by electrical stimulation of the cochlear wall and by cochlear implantation. *Laryngoscope*, 104(6 Pt 1):752–754.

Jolly, C. N., Spelman, F. A. et Clopton, B. M. (1996). Quadrupolar stimulation for cochlear prostheses : modeling and experimental data. *IEEE Trans. Biomed. Eng.*, 43(8):857–865.

Kiefer, J., Gstoettner, W., Baumgartner, W., Pok, S. M., Tillein, J., Ye, Q. et von Ilberg, C. (2004). Conservation of low-frequency hearing in cochlear implantation. *Acta Otolaryngol.*, 124(3):272–280.

Loizou, P. C., Stickney, G., Mishra, L. et Assmann, P. (2003). Comparison of speech processing strategies used in the Clarion implant processor. *Ear Hear.*, 24(1):12–19.

McDermott, H. J., McKay, C. M. et Vandali, A. E. (1992). A new portable sound processor for the university of Melbourne/Nucleus limited multielectrode cochlear implant. *J. Acoust. Soc. Am.*, 91(6):3367–3371.

McKerrow, W. S., Schreiner, C. E., Snyder, R. L., Merzenich, M. M. et Toner, J. G. (1991). Tinnitus suppression by cochlear implants. *Ann. Otol. Rhinol. Laryngol.*, 100(7):552–558.

Miyamoto, R. T., Kirk, K. I., Renshaw, J. et Hussain, D. (1999). Cochlear implantation in auditory neuropathy. *Laryngoscope*, 109(2 Pt 1):181–185.

Peterson, A., Shallop, J., Driscoll, C., Breneman, A., Babb, J., Stoeckel, R. et Fabry, L. (2003). Outcomes of cochlear implantation in children with auditory neuropathy. *J. Am. Acad. Audiol.*, 14(4):188–201.

Ruckenstein, M. J., Hedgepeth, C., Rafter, K. O., Montes, M. L. et Bigelow, D. C. (2001). Tinnitus suppression in patients with cochlear implants. *Otol. Neurotol.*, 22(2):200–204.

Shallop, J. K., Peterson, A., Facer, G. W., Fabry, L. B. et Driscoll, C. L. (2001). Cochlear implants in five cases of auditory neuropathy : postoperative findings and progress. *Laryngoscope*, 111(4 Pt 1):555–562.

Shipp, D. et Nedzelski, J. (1994). Prognostic value of round window psychophysical measurements with adult cochlear implant candidates. *In* Hochmair-Desoyer, I. et Hochmair, E., éditeurs : *Advances in cochlear implants*, pages 79–81. Vienna : Manz.

Sininger, Y. S. et Trautwein, P. (2002). Electrical stimulation of the auditory nerve via cochlear implants in patients with auditory neuropathy. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 189:29–31.

Skinner, M. W., Holden, L. K., Holden, T. A., Dowell, R. C., Seligman, P. M., Brimacombe, J. A. et Beiter, A. L. (1991). Performance of postlinguistically deaf adults with the wearable speech processor (WSP III) and mini speech processor (MSP) of the Nucleus multi-electrode cochlear implant. *Ear Hear.*, 12(1):3–22.

Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzker, G. R., Brunsden, B. et Blocker, B. (2002). CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *J. Assoc. Res. Otolaryngol.*, 3(3):332–350.

Thedinger, B., House, W. F. et Edgerton, B. J. (1985). Cochlear implant for tinnitus. case reports. *Ann. Otol. Rhinol. Laryngol.*, 94(1 Pt 1):10–13.

Trautwein, P. G., Sininger, Y. S. et Nelson, R. (2000). Cochlear implantation of auditory neuropathy. *J. Am. Acad. Audiol.*, 11(6):309–315.

Truy, E. et Lina, G. (2003). Implantation cochléaire de l'enfant. *Arch. Pediatr.*, 10(6):554–564.

Tyler, R. S. (1995). Tinnitus in the profoundly hearing-impaired and the effects of cochlear implants. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 165:25–30.

Tyler, R. S. et Lowder, M. W. (1992). Audiological management and performance of adult cochlear-implant patients. *Ear Nose Throat J.*, 71(3):117–122, 125–128.

von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Sturzebecher, E. et Klinke, R. (1999). Electric-acoustic stimulation of the auditory system. new technology for severe hearing loss. *ORL J. Otorhinolaryngol. Relat. Spec.*, 61(6):334–340.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. et Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, 352(6332):236–238.

Zeng, F. G. (2004). Trends in cochlear implants. *Trends Amplif.*, 8(1):1–34.

Chapitre 8

Réorganisation neurofonctionnelle après une altération sensorielle

Sommaire

8.1 Propriétés de la neuroplasticité 156
8.1.1 Principes généraux de la plasticité 156
8.1.2 Plasticité physiologique et plasticité anatomique 157
8.1.3 Dynamique de la neuroplasticité 157
8.1.4 Mécanismes de la plasticité 159
8.2 Neuroplasticité intramodale 161
8.2.1 Rôle de la plasticité dans le développement auditif 161
8.2.2 Rôle de la plasticité dans le développement visuel 164
8.2.3 Plasticité intramodale après un entraînement sensoriel 166
8.2.4 Plasticité intramodale après altération des entrées sensorielles 169
8.3 Neuroplasticité intermodale après déprivation auditive 171
8.3.1 Effet de la surdité sur la perception visuelle et somesthésique 171
8.3.2 Neuroplasticité visuelle et somesthésique après une surdité 172
8.4 Neuroplasticité intermodale après déprivation visuelle 173
8.4.1 Réorganisation auditive chez l'homme et l'animal aveugle 173
8.4.2 Réorganisation somesthésique chez l'homme et l'animal aveugle 174
Bibliographie 177

Introduction

Le système nerveux animal possède la capacité naturelle de modifier son organisation. De telles modifications interviennent tout au long du développement et de la maturation normale de l'organisme, accompagnent l'acquisition de nouvelles aptitudes cognitives et perceptives chez l'organisme jeune et adulte, et surviennent à la suite de l'endommagement accidentel du système nerveux et tout particulièrement des récepteurs sensoriels. Cette propriété remarquable du système nerveux est appelée *plasticité corticale*, ou *neuroplasticité*.

L'étude de la plasticité corticale a rapidement progressé ces dernières années, montrant que l'organisation neuronale du cerveau jeune ou adulte est très fortement influencée

par les perceptions en provenance de l'environnement. De nombreuses études ont en particulier montré les effets marquants de la perte d'une modalité perceptive sur le réseau neuronal associé aux autres modalités.

Ces études indiquent non seulement des changements comportementaux dans les modalités intactes mais également une réorganisation des fonctions corticales. Ainsi les aires corticales multimodales présentent une amélioration du traitement des informations en provenance des modalités restantes aussi bien chez l'animal que chez l'homme. Il a été également montré que les cortex sensoriels primaires associés aux modalités perdues pouvaient être recrutés par les autres modalités chez l'homme et l'animal.

Cependant, les rôles respectifs des contraintes intrinsèques au réseau cortical et des informations environnementales dans la détermination de la spécificité fonctionnelle du cortex cérébral restent mal connus. De telles informations sont essentielles pour évaluer et prévoir la réussite des implants sensoriels chez l'être humain.

8.1 Propriétés de la neuroplasticité

8.1.1 Principes généraux de la plasticité

Dans le cadre d'une neuroplasticité induite par une déprivation sensorielle, une distinction a été proposée entre les changements *intramodaux* et *intermodaux* (Röder et Neville, 2003). Les systèmes neuraux spécifiques des modalités non déprivées peuvent se réorganiser après la déprivation d'une autre modalité en raison d'une plus grande dépendance de l'individu vis-à-vis des modalités préservées (modifications intramodales). Par ailleurs, si une partie d'un système sensoriel est déafférentée, les parties restantes de ce système sont susceptibles de subir des modifications. D'un autre côté, des aires cérébrales normalement associées avec une modalité sensorielle complètement déprivée peuvent être innervés par les systèmes sensoriels demeurés intacts (modifications intermodales, ou réafférentation fonctionnelle), ce qui se traduit par une expansion des systèmes sensoriels intacts au-delà de leurs frontières.

Röder et Neville (2003) ont proposé trois mécanismes non exclusifs pour tenter de rendre compte de l'ensemble des formes que peut prendre la réorganisation plastique des réseaux neuraux après une altération ou une déprivation sensorielle : 1° les systèmes sensoriels spécifiques des modalités préservées se réorganisent de façon interne afin d'améliorer leur efficacité de traitement (réorganisation intramodale) ; 2° dans les aires multisensorielles, les sous-réseaux préférentiellement associés à la modalité déprivée répondent également aux stimuli provenant des modalités préservées, et ainsi les cellules multisensorielles peuvent développer une plus grande sensibilité aux stimuli en provenance des modalités intactes ; 3° les aires cérébrales normalement associées avec la modalité déprivée peuvent être recrutées par des afférences en provenance des autres modalités intactes (réorganisation intermodale, ou réallocation fonctionnelle).

De façon très générale, Elbert *et al.* (2001) ont résumé quelques-uns des principes de base sous-tendant les mécanismes de la plasticité corticale :

- l'inutilisation ou la désafférentation (par lésions ou privations sensorielles) entraîne le recrutement des aires corticales inutilisées par des neurones d'aires voisines ;
- une utilisation accrue entraîne une expansion de la représentation corticale associée ;
- des entrées synchrones mènent à une fusion des aires corticales associées à ces entrées ;
- des entrées asynchrones mènent à une ségrégation des aires corticales associées à ces entrées.

8.1.2 Plasticité physiologique et plasticité anatomique

On peut classifier la plasticité neuronale en différentes catégories selon la nature des modifications qu'elle induit dans le cerveau et les mécanismes qui sous-tendent ces modifications. Il faut avant tout faire une distinction entre *plasticité physiologique* et *plasticité anatomique* (Birbaumer et Schmidt, 1991). La première fait référence à des changements dans les propriétés de réponses des neurones et des réseaux de neurones, comme par exemple les seuils de décharge, la potentiation à long-terme ou la dépression à long-terme (Cain, 2001; Teyler, 2001). La seconde fait au contraire référence à des changements dans la structure physique du neurone et des neurites, comme par exemple le nombre et la forme des synapses, la taille des corps cellulaires et des arbres dendritiques.

Il est important de noter que les plasticités physiologique et anatomique ne sont pas indépendantes mais interagissent de façon étroite. Par exemple, lors d'une phase initiale d'apprentissage, on peut observer des modifications des taux de décharge des neurones, en liaison possible avec la désinhibition de synapses muettes ou avec des variations dans la sécrétion de neurotransmetteurs présynaptique. Une transition progressive se fait ensuite vers la phase de consolidation, caractérisée par des modifications structurales qui renforcent les modifications physiologiques initiales (Kolb et Whishaw, 1990). Cette subdivision soulève l'idée d'une classification des changements neuraux en fonction du temps écoulé depuis l'évènement déclencheur, avec une distinction entre les changements rapides (dont l'ordre de grandeur varie entre la seconde et l'heure) et les changements lents (variant entre le jour et le mois). Les changements rapides seraient probablement sous-tendus par le démasquage, la désinhibition ou la potentiation de connexions synaptiques pré-existantes (synapses "muettes", ou plasticité physiologique), tandis que les changements lents impliqueraient probablement des modifications anatomiques supplémentaires comme la ramification collatérale des connexions et des remaniements dans la structure même du neurone, contribuant à la consolidation des changements rapides (Roder et Rosler, 2004). Cette distinction soulève en particulier la question de l'aspect dynamique de la plasticité cérébrale, abordée dans la suite de notre exposé (voir « *Dynamique de la neuroplasticité* », § 8.1.3, p. 157).

8.1.3 Dynamique de la neuroplasticité

La différence physique des changements induits respectivement par les plasticités physiologique et anatomique (voir « *Plasticité physiologique et plasticité anatomique* », § 8.1.2, p. 157) montre une différence d'échelle entre le décours temporel des modifications neurales associées. Ainsi la plasticité dite physiologique est celle qui intervient le plus rapidement, à une échelle inférieure à l'heure, tandis que les modifications plastiques anatomiques se produisent plus lentement, à l'échelle des jours ou même des mois. En élargissant cette échelle temporelle jusqu'à englober la vie entière de l'individu, on peut se demander comment la neuroplasticité intervient dans le développement, depuis l'enfance jusqu'à l'âge adulte. Pour aborder cette problématique nous nous proposons de faire la distinction entre la plasticité supportant la maturation des réseaux neuronaux lors du développement et la plasticité supportant la maintenance de ces réseaux neuronaux lors de l'âge adulte.

Plasticité développementale : période critique et période sensible

Les notions de *période sensible* et de *période critique* sont intimement liées aux phénomènes de plasticité corticale. Knudsen (1988) définit les périodes sensibles comme des périodes de la vie pendant lesquelles des modifications de l'expérience sensorielle (comme une déprivation

sensorielle) peuvent provoquer un développement anormal des réseaux cérébraux associés, et les périodes critiques comme des périodes de la vie pendant lesquelles le développement normal des réseaux cérébraux est dépendant de l'expérience sensorielle.

Période sensible auditive Les données concernant l'existence et l'étendue d'une période critique dans le système auditif des animaux sont principalement sous-tendues par les travaux réalisés dans le domaine de l'implantation cochléaire. Deux axes de recherche essentiels ont été privilégiés dans le cadre de ces travaux, le premier s'intéressant au modèle animal des chats sourds congénitaux, tandis le second s'adressait à l'implantation cochléaire chez l'humain, en particulier chez les enfants.

Ainsi les données recueillies chez des chats sourds congénitaux implantés cochléaires démontrent que les capacités de réorganisation plastique sont plus importantes chez les jeunes chats que les chats adultes (Kral *et al.*, 2002). Elles mettent également en évidence l'existence d'une période sensible s'étendant jusqu'à cinq mois après la naissance.

Les données issues de l'implantation cochléaire chez l'humain démontrent quant à elles l'existence d'une période sensible auditive dont la durée se rapproche de la période sensible d'acquisition du langage, comprise entre deux et quatre ans (voir *« Période sensible pour l'implantation cochléaire »*, § 9.3.3, p. 199).

Période sensible visuelle Les données issues de la déprivation visuelle chez les animaux démontrent l'existence d'une période critique visuelle dont l'étendue serait limitée aux premières semaines ou aux premiers mois de la vie chez les chats et les singes (Hubel et Wiesel, 1970; Von Noorden, 1972; Regal *et al.*, 1976; Crair *et al.*, 1998).

L'amblyopie fournit un modèle intéressant pour tester la durée de la période critique visuelle chez l'humain. Lorsque les informations visuelles en provenance des deux yeux sont incongruentes pendant l'enfance, par exemple en cas de strabisme ou de cataracte, on observe une diminution de la vision pour les deux yeux, même si l'un d'entre eux ne présente aucune atteinte décelable (comme dans le cas du strabisme ou de la cataracte monoculaire) : c'est le phénomène d'amblyopie (Keech et Kutschke, 1995; Kiorpes *et al.*, 1998). De plus, dans le cas d'une cataracte congénitale (mono- ou binoculaire), on observe un déficit dans la perception visuelle de la profondeur (Maurer *et al.*, 1989), accompagné par des affections des fonctions visuelles complexes qui se développent lors des premiers mois de la vie (Maurer et Lewis, 2001), comme l'orientation visuelle (Goldberg *et al.*, 2001) et la perception des visages (Le Grand *et al.*, 2001). L'étude du modèle de l'amblyopie a ainsi permis de démontrer que le cortex visuel humain est caractérisé par une période critique probablement située entre 10 et 20 ans (Vaegan et Taylor, 1979; Epelbaum *et al.*, 1993).

L'étude de l'étendue de la période critique visuelle chez l'humain est également rendue possible par la comparaison des processus neurofonctionnels mis en jeu chez les aveugles précoces et tardifs. En particulier, des variations de l'activation des aires visuelles lors de la lecture Braille en fonction de l'âge de la cécité ont été observées dans plusieurs études (Cohen *et al.*, 1997; Buchel *et al.*, 1998; Cohen *et al.*, 1999). En cohérence avec ces résultats, une stimulation magnétique transcrânienne des aires occipitales n'a aucun effet sur les capacités de discrimination tactile des aveugles tardifs, tandis qu'on observe des modifications comportementales dans le cas des aveugles précoces. L'ensemble de ces études suggère que la fenêtre temporelle de plasticité intermodale chez les humains aveugles est limitée à l'enfance et que les processus de réorganisation neurofonctionnelle peuvent jouer un rôle compensatoire si la cécité intervient avant le début de la puberté, aux alentours de 14 ans.

Plasticité adulte

Des expériences lésionnelles chez l'animal adulte ont montré que des neurones ayant des champs récepteurs correspondant à un point particulier sur la rétine développent, après une lésion artificielle de la rétine en ce point, des champs récepteurs correspondant aux zones rétiennes adjacentes au point original (Kaas *et al.*, 1990). Cela démontre la capacité du système visuel adulte à se remodeler en fonction des informations sensorielles afférentes. Des résultats similaires ont été observés pour les autres systèmes senoriels : un apprentissage perceptif et des lésions périphériques locales chez des individus adultes induisent des réorganisations des systèmes somatosensoriel, auditif et moteur (Kaas, 2000, 2001; Fahle et Poggio, 2002).

Toutefois les mécanismes sous-tendant les plasticité adulte et développementale ne sont pas identiques. Par exemple, chez l'animal adulte, une lésion monoculaire de la rétine ne provoque pas de réorganisation des champs récepteurs correspondants, dans la mesure où la préservation de l'autre rétine permet d'assurer la pérennité de ces champs récepteurs. Dans le cas de l'adulte, seule une lésion rétinienne binoculaire au même point provoque la réorganisation des champs récepteurs associés. Au contraire, une lésion rétinienne monoculaire pendant la période critique du développement provoque la mise en place de deux champs récepteurs pour le même neurone, un pour chaque oeil, ce qui peut provoquer au niveau perceptif un dédoublement éventuellement accompagné d'un brouillage de l'image (Chino *et al.*, 2001). De même, des animaux vivant dans un environnement enrichi présentent des modifications neuroanatomiques globalement similaires, avec toutefois des différences quantitatives et qualitatives selon la période de la vie (enfance ou âge adulte) pendant laquelle ils sont sur-stimulés (voir références dans « *Plasticité intramodale après un entraînement sensoriel* », § 8.2.3, p. 166). Cela démontre que, même si les plasticités développementales et adulte partagent de nombreuses caractéristiques communes, elles ne sont pas complètement similaires.

8.1.4 Mécanismes de la plasticité

Le premier des mécanismes à avoir été suggéré pour expliquer le phénomène de la plasticité corticale est le mécanisme hebbien, dont le principe de base est le renforcement des connections entre les cellules nerveuses activées de façon synchrone. Bien que la capacité d'altérer les poids synaptiques par les processus de potentiation à long-terme (LTP : long-term potentiation) et de dépression à long terme (LTD : long-term depression) ait été étudiée de façon plus extensive dans l'hippocampe, le néocortex présente également une telle plasticité, même dans les organismes maturs.

On suppose généralement que les mécanismes de plasticité synaptique, en particulier la potentiation à long-terme (LTP) (Bliss et Lomo, 1973) et la dépression à long-terme (LTD) (Ito, 1986), sont co-responsables des phénomènes d'apprentissage perceptif et des réorganisations consécutives à une altération des entrées sensorielles. La LTP au niveau des synapses thalamo-corticales (couches II, III et IV) peut être provoquée plus facilement chez les jeunes animaux (Crair et Malenka, 1995; Sermasi *et al.*, 1999). La LTD dans le cortex visuel peut être également provoquée plus facilement chez les jeunes animaux, par ailleurs la période de déclenchement de la LTD est plus longue que la période de déclenchement de la LTP (Rittenhouse *et al.*, 1999). Les effets de la LTP et de la LTD au niveau des synapses thalamo-corticales expliquent par exemple le développement de l'amblyopie chez les chats déprivés monoculairement (Kirkwood *et al.*, 1996).

La sensibilité différente du cortex aux inductions de la LTP et la LTD est basée sur des modifications des récepteurs au glutamate et de leur composition pendant le développement. Dans la couche IV du cortex visuel, les récepteurs NMDA présents à la naissance sont remplacés

par des récepteurs AMPA plus tard au cours du développement (Kaczmarek *et al.*, 1997). Les récepteurs NMDA situés dans les autres couches corticales subissent également des modifications moléculaires durant le développement (Hickmott et Constantine-Paton, 1997; Quinlan *et al.*, 1999a,b; Tang *et al.*, 1999). Un réarrangement de leurs sous-unités conduit à des modifications de leur capacité réceptrice, ce qui conduit à une diminution des entrées d'ions calcium après une liaison agoniste, ce qui explique la réduction de la plasticité. Ainsi des changements dans les sous-unités NMDA similaires à ceux observés dans le cortex visuel ont été décrits dans le noyau cochléaire (Forster et Illing, 1998) et peuvent intervenir dans la régulation et la stabilisation synaptique par l'intermédiaire de l'activité sensorielle (Hickmott et Constantine-Paton, 1997; Segal *et al.*, 2000; Colonnese et Constantine-Paton, 2001)

Les mécanismes d'inhibition synaptique semblent par ailleurs être liés à la durée de la période sensible pendant le développement sensoriel : ainsi on observe une augmentation de la proportion de cellules GABA-ergiques vers la fin de la période sensible dans le cortex visuel (Gao et Pallas, 1999; Feldman, 2000). Dans le cortex auditif, le décours temporel de la quantité d'unités GABA-ergiques est similaire à celui du cortex visuel (Gao et Pallas, 1999). L'inhibition a été suggérée comme l'un des facteurs déclenchant la fin de la période sensible (Fagiolini et Hensch, 2000). En particulier, les périodes sensibles ont pu être prolongées grâce à une réduction de l'inhibition lors d'une déprivation sensorielle partielle ou complète, aussi bien pour la modalité visuelle (Quinlan *et al.*, 1999a,b) que pour la modalité auditive (Rajan et Irvine, 1998) ou somesthésique (Fuchs et Salazar, 1998). Par ailleurs, chez le rat à la naissance, on peut remarquer une fonction inhibitrice particulièrement éloignée de celle observée chez les adultes, dans la mesure où les synapses GABA-ergiques de la couche I du cortex jouent un rôle excitateur (Dammerman *et al.*, 2000). Ainsi, de nombreux changements dans la fonction inhibitrice prennent place lors du développement, ces changements pouvant jouer un role plausible dans l'apparition et la disparition d'une période sensible.

D'un autre côté, de nombreuses études démontrent que la plasticité synaptique peut être considérablement modulée par différents facteurs, et ce même dans le cortex adulte. Dans le système auditif, les recherches se sont concentrées sur la modulation cholinergique (Bakin et Weinberger, 1996; Kilgard et Merzenich, 1998; Kirkwood, 2000) ; en effet le système cholinergique semble être l'un des facteurs intervenant dans la plasticité synaptique chez l'adulte. Il a également été suggéré que ce système puisse transformer le mode de transfert d'information de thalamo-cortical en cortico-cortical dans le cortex auditif (Hsieh *et al.*, 2000). Le système cholinergique pourrait aussi intervenir pendant la période sensible (Aramakis *et al.*, 2000), cette hypothèse étant justifiée par la croissance de ce système lors des périodes sensibles du cortex visuel et du cortex auditif (Robertson, 1987; Robertson *et al.*, 1991).

Les facteurs de croissance des nerfs et les facteurs neurotrophiques cérébraux sont aussi d'une importance cruciale pour le développement cortical. Ces facteurs ont un effet sur la durée de la période sensible chez les chats et les rats (Galuske *et al.*, 1999; Pizzorusso *et al.*, 1999; Sermasi *et al.*, 1999). Ils interviennent dans le développement postnatal dépendant des stimulations extérieures, leur production et leurs effets sur la plasticité synaptique et la croissance dendritique étant tous dépendants de l'activité sensorielle (Boulanger et Poo, 1999; Caleo *et al.*, 1999). Ainsi les facteurs neurotrophiques cérébraux subissent une régulation ralentissante après l'ouverture des yeux chez le rat, cette régulation étant retardée chez les rats élevés dans l'obscurité (Capsoni *et al.*, 1999).

8.2 Neuroplasticité intramodale

Une aire corticale est caractérisée par des entrées spécifiques, des réseaux de traitement de l'information, et des sorties, le tout permettant à l'aire de jouer un rôle spécifique dans le comportement. Les approches traditionnelles cherchant à comprendre l'influence de l'environnement sur le développement des réseaux de connexions corticales ont mis en jeu des paradigmes d'altération sensorielle artificielle (Sur *et al.*, 1999; Angelucci *et al.*, 2000). Sur la base de ces approches, nous aborderons dans la suite l'influence des entrées auditives et visuelles sur le développement des réseaux corticaux correspondants.

Une caractéristique commune aux différents systèmes sensoriels est l'existence de projections topographiques depuis la sortie des récepteurs sensoriels jusqu'au système nerveux central. Chez l'humain et chez l'animal on observe ainsi des cartes somatotopiques, rétinotopiques et tonotopiques (ou cochléotopiques) correspondant respectivement aux modalités somesthésique, visuelle et auditive (Syka, 2002). Un usage extensif ou une déafférentation d'une sous-région de l'épithélium sensoriel est susceptible d'induire un remodelage de ces cartes topographiques, cette réorganisation étant sous-tendue par des changements physiologiques et morphologiques des synapses situées dans les relais nerveux de la sensorialité. En particulier, ces changements peuvent inclure une modification des poids synaptiques associée avec une potentiation ou une dépression de long-terme (Teyler, 2001), des variations dans le nombre d'épines dendritiques, dans la morphologie des terminaisons postsynaptiques et dans la croissance des connexions horizontales, des arbres dendritiques et des corps cellulaires (Kolb et Whishaw, 1998).

Après un apprentissage perceptif, on peut observer un élargissement des réseaux neuronaux ou une amélioration de l'efficacité des algorithmes de traitement des informations sensorielles (Gilbert *et al.*, 2001), tandis que dans le cas de lésions périphériques on peut observer un recrutement de la zone déafférentée par les réseaux neuronaux voisins se traduisant par un élargissement des réprésentations sensorielles adjacentes (Buonomano et Merzenich, 1998; Kaas, 2000; Gilbert *et al.*, 2001; Elbert *et al.*, 2001). Les élargissements des représentations sensorielles rencontrés de façon caractéristique dans ces différents cas ont à l'origine été désignés par le terme d'« expansion représentationnelle » (*map expansion*) par Grafman (2000). Les phénomènes de réorganisation intramodale consécutifs à une modification de la stimulation sensorielle ont par ailleurs été observés à différents niveaux corticaux et sous-corticaux, les modifications associées à ces différents étages de traitement des informations sensorielles étant fortement susceptibles d'interagir entre elles (Ergenzinger *et al.*, 1998; Jones et Pons, 1998; Kaas, 1999).

8.2.1 Rôle de la plasticité dans le développement auditif

Chez l'humain, l'activation du système auditif, déterminée par l'apparition de réflexes auditifs nerveux dans le tronc cérébral, commence entre la 26^{e} et la 28^{e} semaine de gestation, et le développement de la fonction auditive continue ensuite pendant plusieurs années. Dans certains modèles animaux utilisés en recherche expérimentale, on observe également un développement précoce de la fonction auditive, comme chez les cobayes et les chinchillas, tandis que pour d'autres espèces on observe un développement altricial de l'audition (après la naissance), comme chez les rats, les souris, les furets, les gerbilles et les chats (Syka, 2002).

Les conséquences fonctionnelles d'une déprivation auditive chez l'animal ne peuvent pas être étudiées de façon aussi simple que pour une déprivation visuelle ; en effet une simple suture des canaux auditifs ne suffit pas à bloquer complètement l'audition. Le décalage des seuils d'audition à la suite d'une simple suture dépasse rarement les 40 dB, par ailleurs l'audition par conduction osseuse n'est pas du tout affectée (Kral *et al.*, 2006). En conséquence, des manipula-

tions mécaniques effectuées sur le méat auditif externe ne peuvent pas supprimer l'audition. De même, la destruction chirurgicale de l'oreille moyenne ne suffit pas, pour la simple raison qu'elle n'affecte pas non plus la conduction osseuse. Ainsi les sons en provenance du corps ainsi que les propres vocalisations ne sont pas atténués par une telle intervention. Seules les destructions cochléaires partielles ou totales permettent de contrôler complètement la déprivation auditive ; toutefois ces destructions sont irréversibles, contrairement aux sutures des paupières dans le cas de la vision.

Déprivation auditive néonatale

De nombreuses études portant sur la neuroplasticité développementale de l'audition se sont focalisées sur l'effet d'une déprivation auditive sur le système auditif immature. Ces travaux ont montré que l'ablation chirurgicale d'une cochlée chez les mammifères néonataux provoque une réduction et une perte des neurones dans le noyau cochléaire (Hashisaki et Rubel, 1989; Moore, 1990; Tierney *et al.*, 1997) et dans le complexe olivaire supérieur (Moore, 1992; Pasic *et al.*, 1994), la formation de nouvelles connexions entre le noyau cochléaire du côté intact et des structures variées dans le tronc cérébral (Kitzes *et al.*, 1995; Russell et Moore, 1995) et le mésencéphale (Nordeen *et al.*, 1983; Moore et Kitzes, 1985; Moore *et al.*, 1994), ainsi qu'une augmentation ipsilatérale de la sensibilité des neurones du colliculus inférieur (Kitzes, 1984; Kitzes et Semple, 1985; Moore *et al.*, 1993) et du cortex auditif primaire (Reale *et al.*, 1987) à une stimulation acoustique de l'oreille intacte. Toutes ces expériences ont mis en jeu des espèces altricials comme le chat, le furet ou la gerbille, chez lesquelles le développement de l'audition comme celui des autres sens se poursuit après la naissance.

Les enregistrements électrophysiologiques réalisés chez le chat adulte normal montrent qu'une proportion réduite des neurones du cortex auditif primaire (aux alentours de 65%) est excitée par un son présenté à l'oreille ipsilatérale, avec des seuils d'excitabilité significativement plus importants pour l'oreille ipsilatérale que pour l'oreille contralatérale. Chez le chat adulte ayant subi une ablation néonatale unilatérale de la cochlée, on observe dans le cortex auditif primaire contralatéral une représentation tonotopique normale ; par ailleurs, une stimulation acoustique de l'oreille préservée montre des seuils d'excitabilité comparables dans le cortex auditif primaire ipsi- et contralatéral (Reale *et al.*, 1987). En conformité avec ces résultats, on observe chez les lapins ayant subi une ablation cochléaire néonatale unilatérale une modification de l'orientation et une augmentation de la longueur des dendrites dans le cortex auditif contralatéral à l'oreille opérée (McMullen *et al.*, 1988). Chez la gerbille ayant subi une ablation cochléaire néonatale unilatérale, le noyau cochléaire contralatéral à l'oreille opérée présentait une augmentation du nombre de cellules projetant sur le colliculus inférieur ipsilatéral en parallèle avec une diminution du nombre de cellules projetant sur le colliculus inférieur contralatéral, en comparaison avec des animaux contrôles (Nordeen *et al.*, 1983; Moore et Kitzes, 1985; Kitzes et Semple, 1985)

En parallèle avec les modifications structurales des connexions entre les noyaux cochléaires et le colliculus inférieur du côté de l'oreille préservée, ces études ont montré une augmentation de l'influence excitatrice de l'oreille préservée sur le colliculus inférieur ipsilatéral, avec une augmentation de l'amplitude des réponses, une augmentation de l'échelle dynamique de réponse et une diminution des seuils d'excitatibilité, aussi bien en enregistrements unitaires (Kitzes et Semple, 1985) qu'en enregistrements multi-unitaires (Nordeen *et al.*, 1983). Les réponses du colliculus inférieur à une stimulation auditive unilatérale chez les animaux opérés ressemblaient aux réponses à une stimulation auditive contralatérale chez les animaux sains. Des études chez les chats ayant subi une ablation cochléaire néonatale unilatérale montrent que la réorganisation fonctionnelle des voies auditives projetant sur le colliculus inférieur n'était pas accompagnée de modifications

de la densité synaptique du colliculus inférieur, de telles modifications n'étant observées qu'en cas d'ablation cochléaire néonatale bilatérale (Hardie *et al.*, 1998). D'autre études chez les gerbilles ont montré des modifications structurales du complexe olivaire supérieur après une ablation cochléaire néonatale unilatérale (Kitzes *et al.*, 1995). En particulier, les projections du noyau cochléaire ventral vers les noyaux thalamiques supérieurs, qui sont normalement dirigées vers le côté contralatéral, étaient au contraire dirigées bilatéralement vers les noyaux du complexe olivaire supérieur (olive supérieure médiale et latérale, noyau moyen du corps trapézoïdal) après ablation cochléaire néonatale unilatérale, et ceci dans un délai de 24h.

Déprivation auditive à l'âge adulte

Un important volume de données expérimentales montre une réorganisation des structures auditives thalamiques et corticales après une déprivation auditive chez l'animal adulte. Ainsi une déprivation auditive par ablation chirurgicale ou par administration d'un ototoxique provoque dans un délai de 48h une réduction de près de 25% des grandes structures cellulaires sphériques du noyau cochléaire ventral chez la gerbille adulte, ce qui semble indiquer l'importante cruciale de l'activité électrique du nerf auditif dans la régulation transneuronale des tailles cellulaires (Pasic et Rubel, 1989). D'autres études lésionnelles similaires montrent une diminution de la taille des cellules dans les noyaux cochléaires déafférentés (Powell et Erulkar, 1962) et dans le noyau médian du corps trapézoïdal contralatéral (Jean-Baptiste et Morest, 1975).

Les résultats concernant la réorganisation des relais auditifs dans le colliculus inférieur après une déprivation auditive chez l'animal adulte sont plus hétérogènes suivant les espèces. Chez la gerbille adulte, une ablation cochléaire unilatérale ne modifie pas la proportion des cellules du colliculus inférieur excitées par l'oreille ipsilatérale (Nordeen *et al.*, 1983). De larges lésions du noyau cochléaire chez la gerbille adulte ne semblent pas provoquer de modification des réponses unitaires du colliculus inférieur à une stimulation contralatérale (dans l'oreille intacte) (Moore et Kitzes, 1986). On obtient des observations différentes chez le furet adulte ayant subi une ablation cochléaire unilatérale, où l'on remarque dans le colliculus inférieur situé du même côté que l'oreille préservée une proportion anormalement élevée de neurones répondant à une stimulation acoustique de cette oreille (Moore et Kowalchuk, 1988). Des études plus poussées montrent une réorganisation rapide du cortex auditif du furet après ablation cochléaire unilatérale, la représentation tonotopique de l'oreille préservée subissant une expansion dans le cortex auditif primaire ipsilatéral (Moore *et al.*, 1997). Par ailleurs, le cortex auditif primaire situé du côté préservé présente une augmentation de la proportion des cellules répondant à une stimulation ipsilatérale en comparaison avec la proportion rencontrée chez les furets non déprivés (approximativement égale à 50%).

En complément avec l'observation de modifications parfois rapides du système auditif central à la suite d'une ablation cochléaire chez l'animal adulte, un aspect dynamique de plus long terme a été observé, avec une augmentation progressive au cours du temps du nombre de sites du colliculus inférieur répondant à une stimulation de l'oreille ipsilatérale, en parallèle avec une diminution proggressive des seuils d'excitabilité correspondants (Moore *et al.*, 1997). Ces résultats concordent avec des études chez le cobaye adulte par administration d'ototoxique provoquant une destruction quasi-complète des cellules ciliées, avec une diminution progressive des seuils d'excitabilité des potentiels évoqués ipsilatéraux à l'oreille préservée ainsi qu'une augmentation progressive de l'amplitude de ces potentiels dans un délai de plusieurs jours à plusieurs semaines après la destruction des cellules ciliées (Popelar *et al.*, 1994).

8.2.2 Rôle de la plasticité dans le développement visuel

Développement des colonnes de dominance oculaire

Les paradigmes expérimentaux s'intéressant au développement et à la plasticité des colonnes de dominance oculaire dans le cortex visuel primaire ont permis d'importantes avancées dans la compréhension des mécanismes de formation et de régulation des réseaux neuronaux thalamocorticaux par les entrées sensorielles visuelles (Katz et Shatz, 1996). Les colonnes de dominance oculaire sont issues d'une ségrégation des informations visuelles en provenance des deux yeux, relayée par une répartition en couches monoculaires dans le noyau genouillé latéral (LGN). De nombreuses études démontrent que des paradigmes de déprivation monoculaire, comme la suturation d'un oeil lors d'une période critique précoce dans le développement, réduisent la proportion de neurones associés à l'oeil suturé (Wiesel et Hubel, 1965) ainsi que la taille des colonnes de dominance oculaire associée à l'oeil suturé (Shatz et Stryker, 1978; Antonini et Stryker, 1993). Ces expériences démontrent une plasticité de la dominance oculaire où l'oeil resté ouvert domine le cortex à la fois physiologiquement et anatomiquement, probablement au travers d'un mécanisme dynamique de compétition entre les axones en provenance des deux yeux pour le territoire cortical et la liaison synaptique avec les cellules cibles.

Toutefois, de telles études ne permettent pas de répondre à la problématique de la similarité ou non des mécanismes impliqués dans la formation puis dans la maintenance des colonnes de dominance oculaire. En réalité la formation des colonnes de dominance oculaire peut intervenir très tôt dans le développement, même pendant la période où les axones en provenance du LGN croissent dans la couche IV du cortex visuel primaire, si bien que les axones en provenance de chaque couche thalamique monoculaire projettent dans des zones ne se chevauchant pas (Crowley et Katz, 2000; Crair *et al.*, 2001). Ces observations sont renforcées par la découverte d'un développement des colonnes de dominance oculaire en l'absence d'entrées rétinales, dans la mesure où la déprivation visuelle ne dégrade pas immédiatement les zones morcelées des terminaisons géniculocorticales (Crowley et Katz, 1999). De la même façon, une déprivation monoculaire très précoce n'a pas d'effet immédiat sur l'organisation binoculaire du cortex visuel primaire, étant ensuite suivie par une réduction des réponses provenant de l'oeil fermé (Crair *et al.*, 1998). Cet ensemble d'expériences semble indiquer que la formation initiale des colonnes de dominance oculaire semble suivre un programme développemental intrinsèque, une influence de l'activité sensorielle sur cette formation initiale étant moins plausible (Sur, 2004). Nous montrerons dans la suite de notre exposé que cette hypothèse est largement contestable.

Développement des colonnes d'orientation

Une deuxième approche permettant de comprendre le rôle de l'activité sensorielle dans le développement des réseaux corticaux visuels fait intervenir la sélectivité à l'orientation. Les neurones sélectifs à l'orientation et les colonnes d'orientation du cortex visuel primaire font intervenir un réseau plus complexe que les terminaisons axonales thalamocorticales dans les régions monoculaires. L'initiation de la sélectivité à l'orientation dans V1 provient d'entrées du LGN alignées selon l'axe d'orientation (Hubel et Wiesel, 1968; Ferster et Miller, 2000) ; par ailleurs de nombreuses observations expérimentales montrent que les entrées thalamiques sont amplifiées par des connexions intracorticales locales qui affinent la réponse à l'orientation et contribuent à contrôler le gain en contraste (Somers *et al.*, 1995, 2001). Les cellules sélectives à l'orientation sont organisées en cartes d'orientation, au moins dans les couches superficielles de V1 (Bonhoeffer et Grinvald, 1991). Pour finir, les colonnes de cellules répondant à une orientation donnée sont préférentiellement reliées par des connexions horizontales à longue distance, ces

dernières étant intrinsèques aux couches superficielles de V1 et importantes pour la génération des cartes d'orientation (Gilbert et Wiesel, 1989).

Comme pour la dominance oculaire, le développement de la sélectivité à l'orientation dans V1 a été étudié de façon extensive et sa plasticité examinée au travers de programmes de déprivation visuelle. La sélectivité à l'orientation est présente dans l'aire V1 des singes à la naissance (Wiesel et Hubel, 1974) et chez les chats et furets lors de l'ouverture des yeux, bien que l'expérience visuelle contribue de façon importante au raffinement de cette sélectivité (Sherman et Spear, 1982; Fregnac et Imbert, 1984; Chapman et Stryker, 1993). Une déprivation visuelle contrarie le développement des réponses sélectives à l'orientation sans toutefois complètement l'empêcher (Sherman et Spear, 1982; Fregnac et Imbert, 1984; Chapman et Stryker, 1993). Bien que les connexions horizontales à longue distance soient présentes juste avant l'ouverture des yeux chez les chats et les furets, ces connexions sont raffinées après la mise en place de la vision, et une déprivation visuelle empêche ce raffinement (Callaway et Katz, 1990, 1991; Ruthazer et Stryker, 1996). Les cartes d'orientation de V1 se développent en parallèle avec le raffinement des connexions horizontales, mais la taille et la distribution des zones d'orientation reste stable au cours du développement (Chapman *et al.*, 1996). Une déprivation visuelle n'altère pas la formation de ces cartes, bien qu'elle réduise l'intensité des signaux et la régularité de la distribution dans la carte associée à l'oeil déprivé (Crair *et al.*, 1998). Dans l'aire V2 du chat, une suture des paupières à court terme après la formation de la carte d'orientation provoque une détérioration de cette carte, tandis qu'une désuturation de l'oeil induit une restauration de cette carte (Kim et Bonhoeffer, 1994). De la même façon que dans le cas des colonnes de dominance oculaire, ces résultats expérimentaux semblent indiquer que l'organisation de la sélectivité à l'orientation et des cartes d'orientation de V1 est avant tout dépendante d'un processus initial de connectivité intrinsèque, tout en étant moins dépendante des influences visuelles (Sur, 2004). Les résultats expérimentaux décrits dans le paragraphe suivant ont en réalité apporté un net démenti à cette hypothèse.

Réafférention artificielle des entrées visuelles sur les centres nerveux auditifs

La problématique du lien entre le développement des connexions corticales et les motifs d'activité sensorielle afférents a pu être envisagée sous un éclairage nouveau grâce au modèle expérimental de réafférentation des axones associées à une modalité perceptive sur les structures nerveuses centrales associées à une autre modalité perceptive (Sur *et al.*, 1990; Sur et Leamey, 2001; Sur, 2004). Les premières tentatives dans cette direction ont été envisagées chez le hamster, où on a pu démontrer la possibilité de rediriger les axones en provenance de la rétine vers le thalamus auditif, à condition de déafférenter ce thalamus chez l'animal nouveau-né (Schneider, 1973; Frost, 1982). La possibilité d'une réafférentation sensorielle artificielle a ensuite été envisagée chez le furet, qui présente l'avantage de naître à un stade très immature du développement du système visuel, les axones de la rétine étant en cours de développement vers le thalamus tandis que ceux du thalamus n'ont pas encore innervé les couches corticales (Sur *et al.*, 1988). Sur *et al.* (1988) ont en particulier indiqué une méthode de réafférentation basée sur une ablation chirurgicale de certaines connexions du noyau genouillé médial (MGN).

Dans le modèle du furet réafférenté, on peut observer dans le MGN l'apparition de cellules possédant des champs récepteurs visuels à centres on/off (Roe *et al.*, 1993), ainsi que l'émergence d'une représentation rétinotopique avec une différenciation des zones fovéale, périphérique, ventrale et dorsale de la rétine (Roe *et al.*, 1991). De plus, les axones rétinaux en provenance de chaque oeil, occupant initialement des zones se chevauchant, subissent une ségrégation progressive

en territoires monoculaires similaire à celle existant dans le LGN normal, bien que ces territoires conservent une taille inférieure à ceux que l'on trouve dans le LGN normal (Angelucci *et al.*, 1997).

Par ailleurs, une carte rétinotopique apparaît dans le cortex auditif primaire du furet réafférenté (Roe *et al.*, 1990). Les neurones de A1 répondent moins aux stimuli visuels que les neurones de V1 chez les furets non réafférentés (Roe *et al.*, 1992), toutefois le cortex auditif primaire montre des cellules sélectives à l'orientation et des cartes d'orientation visuelle dont la sélectivité est comparable à celle du cortex visuel primaire normal (Sharma *et al.*, 2000). L'organisation de ces cartes est également similaire à celle de V1 chez le furet normal, avec une sudivision en domaines répondant à une orientation préférentielle, ces domaines s'organisant autour de centres de rotations. Comme pour V1, les connexions à longue distance dans A1 relient entre eux des domaines répondant aux mêmes orientations.

Les différences entre furet normal et furet réafférenté résident dans une plus grande taille des domaines d'orientation dans A1, avec une distribution moins ordonnée de ces domaines et des connexions horizontales sous-jacentes. Dans le cortex visuel normal, ces connexions sont distribuées périodiquement suivant une structure modulaire, avec une élongation le long de l'axe d'orientation du site d'injection, tandis que dans le cortex auditif normal elles sont distribuées sous formes de bandes étendues le long de l'axe isofréquentiel de la représentation tonotopique corticale. Dans le cortex auditif réafférenté, on observe une distribution en modules nettement plus petits que les bandes normalement observables dans A1, ces motifs s'étendant préférentiellement selon l'axe d'orientation du site d'activation (Gao et Pallas, 1999).

Cette réorganisation corticale particulière s'accompagne d'une récupération fonctionnelle montrant que les furets réafférentés perçoivent réellement les stimuli visuels par l'intermédiaire de leur cortex auditif (von Melchner *et al.*, 2000). von Melchner *et al.* (2000) ont montré en particulier que les furets réafférentés répondaient de façon d'autant plus précise à un contraste d'orientation que ce contraste est élevé ; de même Frost *et al.* (2000) ont montré que les furets étaient capables de discriminer des motifs visuels par l'intermédiaire de leurs projections rétinales redirigées artificiellement sur le cortex auditif.

Cet ensemble de résultats expérimentaux vient fortement tempérer l'hypothèse d'une programmation intrinsèque du développement de la connectivité corticale : ces expériences montrent en effet que cette connectivité peut être considérablement dépendante des entrées sensorielles afférentes aux centres nerveux. Une explication possible de ces résultats est que la disposition du cerveau à traiter des informations visuelles ou auditives réside dans une localisation non spécifiée initialement. Ainsi la réafférentation artificielle des entrées rétiniennes vers le thalamus auditif provoquerait un décalage de cette localisation, qui se traduirait finalement par l'interprétation des activations visuelles des centres nerveux auditifs en tant qu'informations de nature visuelle. En allant un peu plus loin, on peut faire l'hypothèse que le rôle fonctionnel des réseaux cérébraux dérive directement de leurs entrées sensorielles (Sur, 2004). Par conséquent, dans le développement normal de l'individu, les mécanismes de plasticité cérébrale fourniraient un processus qui permet d'associer naturellement les réseaux cérébraux avec leurs entrées sensorielles, sans nécessiter une spécification explicite de la connectivité du réseau ; en d'autres termes les réseaux de traitement cortical seraient caractérisés par une propriété d'auto-organisation.

8.2.3 Plasticité intramodale après un entraînement sensoriel

De nombreuses études impliquant les différentes modalités sensorielles ont démontré que même en l'absence de toute déprivation sensorielle on pouvait observer une plasticité des cartes

corticales chez l'humain et l'animal, aussi bien pendant le développement qu'à l'âge adulte. Ces études ont toutes démontré l'effet d'un entraînement perceptif sur la plasticité des représentations sensorielles corticales, que cet entraînement soit associé ou pas avec une tâche de conditionnement (Kaas, 1991; Gilbert, 1995; Cruikshank et Weinberger, 1996; Buonomano et Merzenich, 1998; Sathian, 1998; Braun *et al.*, 2000; Pantev et Lütkenhöner, 2000; Menning *et al.*, 2000; Kaas, 2000; Gilbert *et al.*, 2001). Toutefois la réorganisation observée dans le cas d'un entraînement sensoriel n'est ni aussi poussée ni aussi étendue que dans le cas d'une lésion périphérique (Pons *et al.*, 1991).

Environnement enrichi

Les animaux élevés dans des environnements enrichis présentent des cortex plus épais et plus lourds, des embranchements dendritiques plus complexes, un plus grand nombres d'épines dendritiques et une plus grande densité de cellules gliales (Turner et Greenough, 1985; Birbaumer et Schmidt, 1991; Black *et al.*, 1998). Ils montrent également des capacités d'apprentissage plus rapides, la différence avec les animaux contrôles élevés en environnement appauvri étant d'autant plus marquée que la tâche demandée est complexe (Black *et al.*, 1998). Chez les animaux adultes, les environnements enrichis induisent des modifications neuroanatomiques semblables, avec toutefois une certaine dégradation dans la quantité (Black *et al.*, 1998) et la qualité de ces changements (Kolb et Gibb, 2001) en comparaison avec les animaux plus jeunes.

Plasticité somesthésique

Des singes adultes ont été entraînés au travers d'une tâche tactile où ils devaient réguler le contact d'une ou deux extrêmités de leurs doigts sur un disque rainuré en rotation, ce qui correspondait à une stimulation différentielle intense dans le cadre d'une tâche cognitive impliquant beaucoup d'attention. Les cartes des aires sensorielles somesthésiques comparées avant et après l'entraînement ont montré une extension très importante des représentations corticales associées spécifiquement aux extrêmités des doigts qui restaient en contact avec le disque rainuré lors de l'entraînement (Jenkins *et al.*, 1990). Cet élargissement était en correspondance avec une réduction des champs récepteurs dans ces représentations corticales spécifiques. Ces résultats montrent que le cortex est susceptible de remodeler ses représentations sensorielles de façon dynamique selon un processus dépendant des stimulations extérieures, et ce pendant toute la durée de la vie. Une autre étude chez le singe a montré un doublement des représentations corticales somesthésiques des extrêmités des doigts impliqués dans une tâche difficile de recherche de petits objets, accompagné d'une réduction des champs récepteurs associés (Xerri *et al.*, 1994). Les mêmes auteurs ont également montré un doublement des représentations corticales de la peau des tétons des rats femelles allaitantes en comparaison avec les rats femelles non allaitantes, également accompagné par une réduction des champs récepteurs correspondants (Xerri *et al.*, 1994).

Dans une autre classe importante d'études sur l'effet d'un entraînement somesthésique, des singes ont été entraînés à discriminer des vibrations tactiles de fréquences différentes. Dans ces études, la position de la main était fixée et tous les stimuli étaient appliqués au même endroit de la peau de l'un des doigts (Recanzone *et al.*, 1992a,b,c,d). La conséquence de cet entraînement spécifique a été une extension considérable des représentations corticales somesthésiques associées, cette fois accompagnée par un élargissement des champs récepteurs dans la région de peau ayant subi l'entraînement, ce qui contraste avec les réductions de champs récepteurs observées dans les autres études (Jenkins *et al.*, 1990; Xerri *et al.*, 1994). Cette observation a été interprétée dans le cadre de la plasticité hebbienne, qui prédit qu'une stimulation en provenance de la même

position de la peau aboutit nécessairement à un élargissement des champs récepteurs associés. Au contraire, pour des stimulations présentées de façon incongruente en différentes positions voisines de la peau, comme c'est le cas lorsque les stimuli se déplacent au contact de la peau (Jenkins *et al.*, 1990) ou sont positionnés de façon non précisément contrôlée (Xerri *et al.*, 1994), chaque position de la peau rentre en compétition hebbienne avec les positions voisines, ce qui résulte en un rétrécissement des champs récepteurs à mesure que la surface de peau impliquée s'élargit (Buonomano et Merzenich, 1998).

L'importance de la corrélation temporelle de stimulation a également été testée dans le cas de la somesthésie. Dans l'une des études correpondantes, des singes ont été entraînés dans une tâche tactile où une barre orientée perpendiculairement aux doigts stimulait alternativement les segments distaux et proximaux des trois doigts centraux de la main, cette stimulation étant systématiquement appliquée de façon simultanée aux trois doigts. Alors que les représentations somesthésiques normales de la main sont délimitées pour chaque segment digital et incluent très rarement de multiples doigts, on a observé chez les singes entraînés un bouleversement de la ségrégation normale entre les représentations des doigts : après l'entraînement, environ 40% des cellules représentant les doigts avaient développé des champs récepteurs multiples possédant des latences et des amplitudes de réponse comparables pour deux ou trois des doigts centraux de la main (Wang *et al.*, 1995). Ce mode anormal de représentation somesthésique peut être également interprété dans le cadre de la plasticité hebbienne, qui prédit qu'une stimulation simultanée des doigts va renforcer la corrélation des réponses neurales associées à chacun des doigts.

De telles observations ont été répétées chez l'humain ; ainsi les représentations somatosensorielles des doigts de la main gauche chez des violonistes sont significativement plus étendues que les représentations des doigts de la main droite chez ces mêmes violonistes ou que les représentations des doigts de la main gauche chez des sujets contrôles (Elbert *et al.*, 1995).

Plasticité auditive

Des modifications plastiques de l'organisation du cortex dépendantes de l'expérience sensorielle ont été enregistrées dans les aires auditives primaires du singe après un entraînement de plusieurs semaines sur une tâche de discrimination fréquentielle. Dans cette expérience, en parallèle avec l'amélioration progressive des seuils de discrimination fréquentielle, la cartographie détaillée de l'organisation tonotopique de A1 met en évidence un élargissement important des zones associées aux fréquences utilisées pendant l'entraînement auditif, en comparaison avec des singes non entraînés. Une corrélation entre les performances comportementales et l'étendue de l'élargissement des représentations tonotopiques a par ailleurs été observée (Recanzone *et al.*, 1993).

Un autre type de plasticité corticale dépendante de l'expérience a été décliné en plusieurs variations par Weinberger *et al.* (1993). Dans un protocole de conditionnement classique, un son pur de fréquence donnée est couplé avec un choc électrique aversif. Les courbes de réponse mesurées sur les mêmes sites du cortex auditif avant et après l'entraînement montrent un décalage des fréquences centrales de réponse dans la direction de la fréquence de conditionnement ; par ailleurs ces décalages se maintiennent ensuite pendant quelques semaines et peuvent être supprimés par un conditionnement extinctif (Diamond et Weinberger, 1986; Bakin et Weinberger, 1990; Weinberger *et al.*, 1993; Weinberger, 1995).

De la même façon, plusieurs études chez l'humain ont démontré que les propriétés plastiques du cortex auditif sont dépendantes de l'expérience sensorielle auditive. Des modifications

spécifiques en fréquence des réponses neurales tonotopiques ont été observées dans un protocole d'imagerie TEP lors d'une tâche de conditionnement aversif visant à discriminer les hautes et basses fréquences (Morris *et al.*, 1998). Une plasticité à court terme du cortex auditif a également été observée en MEG : des sujets ont subi pendant trois jours consécutifs des séances de 3h d'écoute de musique ayant subi un filtrage coupe-bande centré sur 1kHz. Immédiatement après l'écoute de cette musique filtrée, la représentation neurale associée à un stimulus-test à 1kHz centré sur la fréquence centrale du filtre coupe-bande a montré une diminution significative en comparaison avec la représentation neurale associée à un stimulus-test à 500Hz. Par ailleurs cette diminution de la représentation à 1kHz a disparu systématiquement entre les différentes sessions d'écoute (Pantev *et al.*, 1999). La plasticité à court terme du cortex auditif a également été démontrée au travers de l'étude des corrélats neurophysiologiques de l'image résiduelle auditive ("Zwicker tone"). Cette image résiduelle peut être déclenchée par la suppression soudaine d'un bruit coupe-bande présenté pendant une certaine durée, qui induit alors la sensation d'un son pur dont la fréquence correspond à la fréquence centrale du filtre coupe-bande. Une activité neuromagnétique soutenue est alors observée dans le cortex auditif supratemporal, en association avec cette sensation auditive fantôme (Hoke *et al.*, 1996). Dans un protocole d'étude par EEG, des sujets ont été entraînés à discriminer deux prononciations légèrement différentes d'un même phonème (/da/), leurs capacités de différenciation étant testées au travers d'un paradigme de MMN ("mismatch-negativity"), c'est-à-dire une réponse corticale enregistrée lors de la présentation d'un stimulus déviant dans une série de stimuli identique. Les résultats ont démontré que l'entraînement auditif pouvait influencer la réponse neurophysiologique du système auditif central à de très faibles différences dans la prononciation de la parole (Kraus *et al.*, 1995).

Une plasticité de long-terme a été également démontrée dans le cortex auditif humain, par exemple dans le cas des musiciens expérimentés. Dans le cas des pianistes professionnels, on a observé en imagerie par MEG un élargissement de près de 25% des moments dipôlaires observés en réponse aux sons de piano, en comparaison avec des sujets contrôles n'ayant jamais joué d'un instrument de musique. Cet élargissement n'a par ailleurs pas été observé pour des sons purs de fréquences et d'intensité comparables aux sons de piano. On a également observé une corrélation entre l'étendue de cet élargissement et l'âge de début de la pratique du piano, sans différence entre les musiciens ayant une oreille absolue et ceux ayant une oreille relative (Pantev *et al.*, 1998).

8.2.4 Plasticité intramodale après altération des entrées sensorielles

De nombreuses études ont démontré que les cartes sensorielles centrales sont soumises à une réorganisation plastique à la suite d'une modification des entrées sensorielles, tout particulièrement lorsque cette dernière intervient lors des premiers stades du développement. Par exemple, après une déafférentation partielle des parties d'un système sensoriel, il est possible d'observer une réorganisation des cartes sensorielles associées (Kaas, 1991; Cruikshank et Weinberger, 1996; Buonomano et Merzenich, 1998; Kaas, 2000). Ainsi, dans le cas de la modalité somesthésique, les neurones possédant un champ récepteur sur ce doigt développent des champs récepteurs sur les doigts adjacents en cas d'amputation du doigt (Merzenich et Jenkins, 1993). La réorganisation des cartes somatosensorielles a également démontrée dans le cas de l'amputation d'un membre chez l'humain (Elbert *et al.*, 1994; Flor *et al.*, 1995; Grusser *et al.*, 2001). Dans le cas de la modalité visuelle, des observations similaires ont été faites après une lésion rétinienne (Kaas *et al.*, 1990; Chino *et al.*, 1992; Eysel *et al.*, 1999). Dans le cas de la modalité auditive, on retrouve encore le même type de phénomène après une lésion cochléaire (Irvine, 2000) et chez les personnes souffrant d'acouphènes (Hoke *et al.*, 1989; Muhlnickel *et al.*, 1998).

Réorganisation auditive après une lésion cochléaire partielle

Chez les animaux, une lésion cochléaire partielle à la naissance est susceptible d'induire des réorganisations des représentations tonotopiques dans le colliculus inférieur ainsi que dans le cortex auditif primaire. Harrison *et al.* (1991) ont induit une perte néonatale des hautes fréquences par injection d'ototoxiques chez des chatons nouveaux-nés, puis analysé les représentations tonotopiques dans le cortex primaire une année après la lésion. Ils ont observé une réorganisation extensive du cortex auditif primaire, où des régions normalement excitées par les hautes fréquences chez les animaux contrôles étaient recrutées par les basses fréquences, en particulier celles correspondant à la frontière de la lésion. Toutefois, l'injection de traceurs rétrogrades dans différentes régions de A1 a révélé une répartition normale des marquages dans la division ventrale du corps genouillé médial (Stanton et Harrison, 2000). Ainsi les chats lésés n'ont pas développé de projections thalamocorticales supplémentaires, ce qui indique que la réorganisation cochléotopique de A1 n'est pas provoquée par une extension anormale des afférences thalamocorticales. Des expériences lésionnelles similaires menées à la naissance chez les chinchillas (un animal achevant son développement auditif de façon précoce) montrent par ailleurs des réorganisations tonotopiques au niveau du colliculus inférieur : dans ce cas on observe dans la partie ventrale des noyaux centraux du colliculus inférieur des régions « monotones », ou isofréquentielles, à l'emplacement où se trouvent normalement les représentations des hautes fréquences (Harrison *et al.*, 1996). Les réponses fréquentielles des neurones dans ces régions monotones étaient pathologiques et pratiquement identiques entre elles, et correspondaient aux fréquences marquant la frontière de la lésion cochléaire.

On observe également des réorganisations tonotopiques des voies nerveuses auditives dans le cas de lésions cochléaires chez l'animal adulte. Ces observations sont nettement intéressantes dans la mesure où elles viennent contredire la conception maintenant dépassée d'une stabilisation définitive des cartes représentationnelles à l'âge adulte. Une étude pionnière dans ce domaine a été publiée par (Robertson et Irvine, 1989) au sujet de cobayes adultes ayant subi des lésions mécaniques restreintes unilatérales de la cochlée. L'étendue de la perte auditive résultante a ensuite été mesurée par enregistrement de l'audiogramme des potentiels d'actions composés puis par examen histologique de la cochlée. La lésion appliquée produisait un affaiblissement de la sensibilité périphérique dans une bande fréquentielle étroite s'étendant de 10 à 20 kHz, en association avec une région de cellules ciliées internes et externes anormales circonscrite à la portion correspondante de la cochlée. L'organisation fréquentielle du cortex auditif contralatéral à la cochlée lésée a été étudiée par implantation de micro-électrodes entre 35 et 81 jours après la lésion. La région du cortex qui aurait normalement été occupée par une représentation ordonnée et continue des fréquences s'étendant de 10 à 20 kHz s'est révélée occupée par des représentations élargies de bandes fréquentielles étroites correspondant aux frontières des lésions cochléaires (9.9–10.8 kHz et 20.6–22.0 kHz). Les réponses des groupes de neurones présents dans ces régions réorganisées étaient marquées, avec des seuils d'excitabilité dans les bandes fréquentielles nouvellement acquises comparables à ceux observés chez les animaux contrôles non lésionnés.

Ces observations ont été reproduites pour d'autres espèces animales. Des lésions cochléaires unilatérales dans les hautes fréquences, avec une bande fréquentielle s'étendant de 18 à 24 kHz, ont été provoquées mécaniquement chez des chats adultes (Rajan *et al.*, 1993). 2 à 11 mois après, l'observation du cortex auditif contralatéral à la cochlée lésée montre que les régions du cortex auditif primaire qui auraient normalement représenté les fréquences touchées par la lésion avaient été recrutées par une représentation élargie des fréquences correspondant aux limites de la lésion. De façon tout à fait similaire, des lésions cochléaires unilatérales des hautes fréquences provoquées par ototoxicité chez les macaques adultes provoquent, 80 à 90 jours après

le traitement, un recrutement des régions de A1 normalement associées aux hautes fréquences par une représentation élargie des fréquences frontières de la lésion (Schwaber *et al.*, 1993).

Rajan *et al.* (1993) ont également étudié l'impact des lésions cochléaires chez les chats adultes au niveau du cortex auditif ipsilatéral à la lésion. Contrairement au cortex auditif contralatéral à la lésion, le cortex ipsilatéral ne présente aucune différence avec celui des animaux contrôles non lésés, et les courbes de réponses fréquentielles sont également comparables avec celles des animaux normaux. Cette observation est à contraster avec les conséquences d'une ablation cochléaire unilatérale complète, qui induit un renforcement de l'excitabilité du cortex ipsilatéral à la lésion par l'oreille préservée. La différence observée entre les réorganisations ipsi- et contralatérales au niveau des projections dans le cortex auditif primaire laisse donc penser que les modifications plastiques de A1 sont provoquées par des modifications des étages sous-corticaux dans le cas d'une lésion cochléaire unilatérale partielle. Des expériences lésionnelles similaires n'ont pas démontré de réorganisation plastique des cartes représentationnelles dans le noyau cochléaire dorsal, tandis que l'observation du colliculus inférieur a montré des signes irréguliers de changements plastiques (Rajan et Irvine, 1998). Des lésions cochléaires unilatérales partielles provoquées chez le hamster par une exposition intense au bruit n'ont pas non plus montré de réorganisation des noyaux cochléaires dorsaux (Kaltenbach *et al.*, 1992). Ces données suggèrent que les voies nerveuses partant des noyaux cochléaires dorsaux pour arriver aux noyaux centraux du colliculus inférieur ne sont pas impliquées dans la réorganisation plastique du colliculus inférieur consécutive à une lésion cochléaire partielle.

Mécanismes de réorganisation des cartes topographiques

Une explication possible pour la réorganisation des cartes corticales sensorielles est basée sur la perte de l'inhibition corticale, tout particulièrement l'inhibition environnante aux zones lésées (Syka, 2002; Roder et Rosler, 2004). Une atteinte partielle des organes récepteurs induit en effet une perte anatomiquement vérifiée des afférences inhibitrices dans les modalités visuelle (Hendry et Jones, 1986) et somesthésique (Warren *et al.*, 1989; Welker *et al.*, 1989; Garraghty *et al.*, 1991), ainsi qu'un démasquage physiologiquement vérifié des afférences normalement silencieuse dans le cortex sensoriel déprivé (Calford et Tweedale, 1988). L'inhibition environnante aux zones lésées est une inhibition issue des régions nerveuses sensorielles adjacentes à celles émettant des signaux excitateurs ; il a été démontré que l'inhibition environnante restreint l'étendue corticale de la réponse physiologique (Syka, 2002).

8.3 Neuroplasticité intermodale après déprivation auditive

Sur la base des données de la littérature, nous décrivons ici les conséquences d'une déprivation auditive sur les modalités visuelle et somesthésique à partir des points de vue fonctionnel et neurologique. Seuls les traitements visuel et somesthésique de bas niveau sont envisagés dans la suite, une discussion détaillée des conséquences de la déprivation auditive sur la compréhension de la parole étant fournie dans « *Effets de la surdité sur la perception de la parole* », sect. 6.4, p. 137.

8.3.1 Effet de la surdité sur la perception visuelle et somesthésique

Chez l'homme, dans le domaine de la vision, on n'observe pas de différences particulières entre les individus sourds et les individus normo-entendants concernant des fonctions visuelles

de base comme la discrimination de la luminosité (Bross, 1979), le seuil temporel de discrimination visuelle (Bross et Sauerwein, 1980; Mills, 1985; Poizner et Tallal, 1987), la sensibilité au contraste (Finney *et al.*, 2001; Stevens et Neville, 2006), la sensibilité à la direction du mouvement (Bosworth et Dobkins, 1999, 2002a,b) ou à la vitesse du mouvement (Brozinsky et Bavelier, 2004).

En cas de surdité chez l'homme, les principaux changements observés dans la fonction visuelle semblent se limiter aux cas des sourds signeurs. Ces changements sont très spécifiques et concernent essentiellement la répartition spatiale de l'attention visuelle (Bavelier *et al.*, 2006). Ainsi, pour de nombreux aspects de la fonction visuelle en champ périphérique, on peut remarquer une nette amélioration des performances perceptives, comme dans le cas de la détection de stimuli périphériques (Loke et Song, 1991), la détection du mouvement périphérique (Neville et Lawson, 1987; Bavelier *et al.*, 2001; Bavelier et Neville, 2002), l'orientation (Parasnis et Samar, 1985; Bosworth et Dobkins, 2002a) et le traitement de distracteurs périphériques (Proksch et Bavelier, 2002; Bosworth et Dobkins, 2002a; Sladen *et al.*, 2005).

Cette amélioration des performances perceptives des sourds signeurs dans leur champ visuel périphérique semble être accompagnée d'une diminution de certains aspects de la perception visuelle dans le champ central, comme la capacité d'attention soutenue (Parasnis *et al.*, 2003) ou le traitement de distracteurs centraux (Proksch et Bavelier, 2002). Toutefois la plupart des autres fonctions visuelles en champ central restent comparables aux personnes normo-entendantes, en particulier la détection de stimuli centraux (Loke et Song, 1991), la recherche visuelle (Stivalet *et al.*, 1998; Rettenbach *et al.*, 1999) et l'orientation (Parasnis et Samar, 1985; Parasnis, 1992).

Enfin, dans le domaine somesthésique, on peut également observer des capacités de détection supranormales chez les personnes sourdes pour les stimuli tactiles déviants, tandis que les seuils de discrimination des fréquences tactiles sont similaires à ceux des normo-entendants (Levänen et Hamdorf, 2001).

8.3.2 Neuroplasticité visuelle et somesthésique après une surdité

Les modifications fonctionnelles des systèmes visuel et somesthésique induites par la déprivation auditive peuvent être sous-tendues par différents mécanismes de réorganisation corticale. Tout d'abord, les aires corticales associatives recevant des entrées en provenance de différentes modalités sensorielles sont susceptibles de s'adapter à l'absence d'informations auditives en se réorganisant de façon à permettre une plus grande sensibilité aux modalités sensorielles préservées, comme la vision ou le toucher. Cette hypothèse est confortée par des données montrant des modifications dans les aires pariétales multisensorielles des personnes sourdes utilisant le langage des signes (Bavelier *et al.*, 2001).

D'autre part, l'existence de projections nerveuses des aires auditives vers les aires visuelles de bas niveau, incluant les aires visuelles primaires (voir *« Intégration dans les aires corticales de bas-niveau »*, § 5.4.2, p. 116), pourrait faire partie des substrats anatomiques permettant la mise en place d'une réorganisation intermodale après la déprivation auditive. En conséquence il est possible qu'en l'absence d'informations auditives afférentes les aires visuelles deviennent plus disponibles pour des modulations intramodales, en particulier au niveau attentionnel, comme le montrent les données comportementales disponibles chez les sujets humains sourds signeurs (Bavelier *et al.*, 2006).

Sur ce point, les différences entre les traitements corticaux visuels chez les individus sourds et normo-entendants viennent conforter les observations comportementales. Parmi les premières études, Neville et Lawson (1987) ont mis en évidence des temps de réaction plus faibles et des potentiels évoqués plus importants chez les personnes sourdes signeuses lors du traitement de

stimuli visuels périphériques, tandis que les résultats étaient comparables chez les deux populations pour des stimuli centraux. Des données ultérieures en imagerie fonctionnelle montrent un recrutement accru des aires visuelles du mouvement (MT/MST) lors d'une tâche de détection dans le champ visuel périphérique, cet accroissement n'étant plus observé lors d'une détection fovéale (Bavelier *et al.*, 2000).

Enfin, le cortex auditif est également susceptible de se réorganiser afin d'assurer d'autres fonctions sensorielles comme la vision. Cette hypothèse est soutenue par plusieurs études comparatives menées chez des personnes normo-entendantes et sourdes, démontrant chez ces dernières une suractivation des aires auditives de bas niveau du STS — situées caudalement au cortex auditif primaire — par des stimuli visuels, signés et tactiles (Levänen *et al.*, 1998; Neville *et al.*, 1998; Nishimura *et al.*, 1999; Petitto *et al.*, 2000; Bavelier *et al.*, 2001; Finney *et al.*, 2003; Fine *et al.*, 2005). Cette réorganisation pourrait relever d'un mécanisme de compétition similaire à celui observé dans les aires associatives : en effet, un nombre croissant d'études met en évidence le traitement par la partie caudale du cortex auditif temporal d'informations non-auditives, en particulier des informations tactiles et différentes informations visuelles comme les mouvements biologiques et les visages (Allison *et al.*, 2000; Schroeder *et al.*, 2003; Brosch *et al.*, 2005) (voir également *« Intégration dans les aires corticales associatives »*, § 5.4.1, p. 114 et *« Intégration dans les aires corticales de bas-niveau »*, § 5.4.2, p. 116). Cette hypothèse est d'autant plus cohérente que plusieurs études démontrent une corrélation positive entre l'étendue de la plasticité intermodale dans les aires auditives et le degré de perte auditive (Lambertz *et al.*, 2005) ainsi que la durée de la surdité (Sadato *et al.*, 2004).

8.4 Neuroplasticité intermodale après déprivation visuelle

8.4.1 Réorganisation auditive chez l'homme et l'animal aveugle

Modèles animaux

L'observation qualitative des chats ayant subi une suture binoculaire des paupières pendant plusieurs mois à partir de la naissance ne montre pas de détérioration du comportement. Des mesures quantitatives du comportement lors de la localisation des sons confirment cette première impression. L'erreur de localisation sonore a été mesurée chez le chat dans une tâche utilisant une présentation aléatoire de l'azimuth de la source, avec présentation d'une récompense en nourriture lorsque le chat se dirigeait vers la source sonore. L'erreur de localisation était significativement plus faible chez les chats visuellement déprivés que chez les chats normaux (Rauschecker et Kniepert, 1994). En particulier, l'amélioration de la localisation était le plus marquée pour des positions spatiales latérales et arrières, cette amélioration n'étant pas acoompagnée d'une détérioration de la localisation pour les sources situés en face de l'animal. Des résultats identiques ont été montrés pour des furets ayant subi une déprivation visuelle à la naissance (King et Semple, 1996; King et Parsons, 1999).

Ces résultats comportementaux sont confortés par des observations neurophysiologiques chez les animaux ayant subi une déprivation visuelle montrant une augmentation des neurites (dendrites et synapses) dans les aires auditives des souris (Gyllensten *et al.*, 1966) et des rats (Ryugo *et al.*, 1975).

Modèle humain

L'étude des performances auditives chez l'homme aveugle est en conformité avec les observations réalisées dans le cas de la déprivation visuelle des animaux, révélant une aptitude accrue

des personnes aveugles à localiser les sources sonores dans l'espace (Rice *et al.*, 1965; Rice, 1970; Muchnik *et al.*, 1991; Lessard *et al.*, 1998; Roder *et al.*, 1999; Weeks *et al.*, 2000). Les premières études, réalisées dans des paradigmes d'écholocation, ont montré une plus forte amélioration de la localisation dans l'espace latéral (Rice *et al.*, 1965; Rice, 1970). Ces observations ont été confirmées et affinées par la suite, révélant chez l'homme aveugle des capacités de discrimination spatiale auditive particulièrement accrues dans l'espace azimutal latéral, sans détérioration de la localisation frontale (Roder *et al.*, 1999), ce qui rejoint les résultats observés chez le chat (Rauschecker et Kniepert, 1994). D'autres études concordent pour montrer une meilleure capacité globale de traitement auditif spatial chez le sujet aveugle (Ashmead *et al.*, 1998; Voss *et al.*, 2004; Yabe et Kaga, 2005).

Les différences comportementales dans le traitement des informations auditives chez les aveugles comprennent également un meilleur traitement des informations spectrales (Doucet *et al.*, 2005), du ton sonore (Witkin *et al.*, 1968; Gougoux *et al.*, 2004; Hamilton *et al.*, 2004), des informations temporelles (Muchnik *et al.*, 1991; Stevens et Weaver, 2005) et de l'intensité sonore (Benedetti et Loeb, 1972). Hamilton *et al.* (2004) ont également mis en évidence une prévalence accrue de l'oreille absolue chez les musiciens aveugles.

Au niveau cortical, plusieurs études ont mis en évidence un traitement par le cortex visuel des informations auditives. Ainsi Weeks *et al.* (2000) ont démontré que les aires visuelles occipitales associatives qui sont impliquées dans la localisation visuelle et la détection visuelle du mouvement s'activent lors d'une tâche de localisation sonore chez des sujets aveugles congénitaux. Une étude récente a démontré que les sujets aveugles précoces activent leur cortex visuel lors du traitement attentionnel de changements dans l'environnement auditif (Kujala *et al.*, 2005). Des observations de ce type ont pu être reproduites pour différentes fonctions auditives, comme la discrimination des sons (Kujala *et al.*, 1995) ou le traitement des sons de parole (Roder *et al.*, 2002).

8.4.2 Réorganisation somesthésique chez l'homme et l'animal aveugle

Modèles animaux

Les chats et les souris ayant subi une déprivation visuelle à la naissance présentent une croissance anormale des vibrisses faciales et une expansion des représentations somesthésiques corticales associées (Rauschecker *et al.*, 1992). En particulier, chez les rats et les souris énucléés à la naissance, l'expansion corticale des zones de projection correspondant aux vibrisses faciales peut atteindre 30% (Rauschecker *et al.*, 1992; Toldi *et al.*, 1994).

Modèle humain

Les recherches comportementales visant à investiguer les aptitudes somesthésiques chez l'humain aveugle ne montrent pas de net consensus sur l'éventualité d'une différence entre les aveugles et les normo-voyants. Plusieurs études n'indiquent aucune différence entre ces deux groupes, au travers de protocoles évaluant le seuil de détection tactile par l'intermédiaire de monofilaments de nylon (Heinrichs et Moorhouse, 1969; Pascual-Leone et Torres, 1993; Sterr *et al.*, 1998), le seuil de détection des vibrations (Heinrichs et Moorhouse, 1969), le seuil de détection des impulsions électriques (Pascual-Leone et Torres, 1993) ainsi que l'acuité spatiale tactile (Grant *et al.*, 2000). D'un autre côté, plusieurs études montrent des différences entre les seuils tactiles des doigts en fonction de leur utilisation usuelle ou pas lors de la lecture en Braille (Sterr *et al.*, 1998), des différences dans le seuil de précision spatiale tactile (Van Boven *et al.*, 2000; Goldreich et Kanics, 2003) et dans diverses tâches passives d'acuité tactile (Stevens *et al.*,

1996), les sujets aveugles présentant dans tous ces cas des aptitudes tactiles plus développées que les sujets sains. Goldreich et Kanics (2003) ont indiqué plusieurs sources de variabilité inhérentes aux protocoles utilisés dans l'évaluation des compétences tactiles des sujets aveugles et sains, qui pourraient éventuellement expliquer la variabilité des résultats obtenus dans les différentes études. Un compromis récent a été proposé entre les deux extrêmes en suggérant que les modifications comportementales en cas de cécité ne seraient pas d'ordre sensoriel mais d'ordre attentionnel, les sujets aveugles développant de plus grandes capacités d'attention aux stimuli sensoriels (Collignon *et al.*, 2006).

Concernant les corrélats neuronaux associés à la représentation somesthésique en cas de déprivation visuelle, une étude en MEG a démontré chez les aveugles pratiquant la lecture Braille des expansions des représentations pour l'index droit, servant à la lecture Braille, comparé avec l'index de l'autre main, inutilisé lors de la lecture Braille, ou comparé avec l'index droit chez des aveugles ne pratiquant pas la lecture Braille. Par ailleurs, les représentations des parties de la main ne servant pas à la lecture Braille se sont révélées moins étendues que pour les mains contrôles (Pascual-Leone et Torres, 1993).

Sadato *et al.* (1996) ont comparé les activations cérébrales des aveugles précoces et des contrôles normo-voyants dans des tâches de lecture en Braille et de discrimination tactile. Les sujets aveugles ont présenté des activations des cortex visuels primaires et associatifs lors de la lecture en Braille, avec une diminution de ces activations pour les autres tâches de discrimination tactile, tandis que les sujets contrôles ont présenté des désactivations des aires corticales visuelles. Par comparaison, les aires occipitales sont restés inactives dans les deux groupes lors de tâches tactiles simples n'impliquant pas d'efforts de discrimination. Ainsi, chez les sujets aveugles, les aires corticales normalement dédiées aux entrées visuelles sont actives lors de tâches de discrimination tactile, ce qui est en accord avec les résultats d'études moins récentes (Wanet-Defalque *et al.*, 1988; Uhl *et al.*, 1991). Une étude longitudinale ultérieure utilisant des tâches de discrimination tactile sans lecture Braille a montré que les sujets normo-voyants activent les aires somesthésiques secondaires tout en désactivant les aires occipitales, tandis que les sujets aveugles activent les aires occipitales, en particulier V1 (Sadato *et al.*, 1998).

Ces observations, répétées dans de nombreuses études (Sadato *et al.*, 1996, 1998; Cohen *et al.*, 1999; Melzer *et al.*, 2001; Sadato *et al.*, 2002) suggèrent que le réseau neuronal normalement utilisé pour la discrimination des formes visuelles permet chez l'aveugle le traitement de l'information spatiale somesthésique. Cette hypothèse est renforcée par des données lésionnelles montrant chez un sujet aveugle la perte de la capacité à lire le Braille à la suite d'une ischémie occipitale (Hamilton *et al.*, 2000). Par ailleurs, le rôle fonctionnel des activations occipitales dans la médiation de la modalité somesthésique a été confirmé par des études en stimulation magnétique transcrânienne (Cohen *et al.*, 1997, 1999; Pascual-Leone et Hamilton, 2001)

Bibliographie

Allison, T., Puce, A. et McCarthy, G. (2000). Social perception from visual cues : role of the STS region. *Trends Cogn. Sci.*, 4(7):267–278.

Angelucci, A., Clasca, F., Bricolo, E., Cramer, K. S. et Sur, M. (1997). Experimentally induced retinal projections to the ferret auditory thalamus : development of clustered eye-specific patterns in a novel target. *J. Neurosci.*, 17(6):2040–2055.

Angelucci, A., Sharma, J. et Sur, M. (2000). The modificability of neocortical connections and function during development. *In* Kaas, J. H., éditeur : *The mutable brain*, pages 351–392. Amsterdam : Harwood Academic.

Antonini, A. et Stryker, M. P. (1993). Rapid remodeling of axonal arbors in the visual cortex. *Science*, 260(5115):1819–1821.

Aramakis, V. B., Hsieh, C. Y., Leslie, F. M. et Metherate, R. (2000). A critical period for nicotine-induced disruption of synaptic development in rat auditory cortex. *J. Neurosci.*, 20(16):6106–6116.

Ashmead, D. H., Wall, R. S., Ebinger, K. A., Eaton, S. B., Snook-Hill, M. M. et Yang, X. (1998). Spatial hearing in children with visual disabilities. *Perception*, 27(1):105–122.

Bakin, J. S. et Weinberger, N. M. (1990). Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. *Brain Res.*, 536(1-2):271–286.

Bakin, J. S. et Weinberger, N. M. (1996). Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proc. Natl. Acad. Sci. U.S.A.*, 93(20):11219–11224.

Bavelier, D., Brozinsky, C., Tomann, A., Mitchell, T., Neville, H. et Liu, G. (2001). Impact of early deafness and early exposure to sign language on the cerebral organization for motion processing. *J. Neurosci.*, 21(22):8931–8942.

Bavelier, D., Dye, M. W. et Hauser, P. C. (2006). Do deaf individuals see better ? *Trends Cogn. Sci.*, 10(11):512–518.

Bavelier, D. et Neville, H. J. (2002). Cross-modal plasticity : where and how ? *Nat. Rev. Neurosci.*, 3(6):443–452.

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G. et Neville, H. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *J. Neurosci.*, 20(17):93.

Benedetti, M. et Loeb, M. (1972). A comparison of auditory monitoring performance in blind subjects and that of sighted subjects in light and dark. *Percept. Psychophys.*, 11:10.

Birbaumer, N. et Schmidt, R. F. (1991). *Biologische Psychologie*. Berlin : Springer.

Black, J. E., Jones, T. A., Nelson, C. A. et Greenough, W. T. (1998). Neuroplasticity and the developing brain. *In* Alessi, N., Coyle, J. T., Harrison, S. I. et Eth, S., éditeurs : *The handbook of child and adolescent psychiatry*, pages 31–53. New York : Wiley.

Bliss, T. V. et Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, 232(2):331–356.

Bonhoeffer, T. et Grinvald, A. (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, 353(6343):429–431.

Bosworth, R. G. et Dobkins, K. R. (1999). Left hemisphere dominance for motion processing in deaf signers. *Psychol. Sci.*, 10:256.

Bosworth, R. G. et Dobkins, K. R. (2002a). The effects of spatial attention on motion processing in deaf signers, hearing signers, and hearing nonsigners. *Brain Cogn.*, 49(1):152–169.

Bosworth, R. G. et Dobkins, K. R. (2002b). Visual field asymmetries for motion processing in deaf and hearing signers. *Brain Cogn.*, 49(1):170–181.

Boulanger, L. et Poo, M. M. (1999). Presynaptic depolarization facilitates neurotrophin-induced synaptic potentiation. *Nat. Neurosci.*, 2(4):346–351.

Braun, C., Wilms, A., Schweizer, R., Godde, B., Preissl, H. et Birbaumer, N. (2000). Activity patterns of human somatosensory cortex adapt dynamically to stimulus properties. *Neuroreport*, 11(13):2977–2980.

Brosch, M., Selezneva, E. et Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.*, 25(29):6797–6806.

Bross, M. (1979). Residual sensory capacities of the deaf : a signal detection analysis of a visual discrimination task. *Percept. Mot. Skills*, 48(1):187–194.

Bross, M. et Sauerwein, H. (1980). Signal detection analysis of visual flicker in deaf and hearing individuals. *Percept. Mot. Skills*, 51(3 Pt 1):839–843.

Brozinsky, C. J. et Bavelier, D. (2004). Motion velocity thresholds in deaf signers : changes in lateralization but not in overall sensitivity. *Brain Res. Cogn. Brain Res.*, 21(1):1–10.

Buchel, C., Price, C., Frackowiak, R. S. et Friston, K. (1998). Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain*, 121:3.

Buonomano, D. V. et Merzenich, M. M. (1998). Cortical plasticity : from synapses to maps. *Annu. Rev. Neurosci.*, 21:149–186.

Cain, D. P. (2001). Synaptic models of neuroplasticity : What is LTP ? *In* Shaw, C. A. et McEachern, J. C., éditeurs : *Toward a theory of neuroplasticity*, pages 118–129. Hove, England : Psychology Press.

Caleo, M., Lodovichi, C. et Maffei, L. (1999). Effects of nerve growth factor on visual cortical plasticity require afferent electrical activity. *Eur. J. Neurosci.*, 11(8):2979–2984.

Calford, M. B. et Tweedale, R. (1988). Immediate and chronic changes in responses of somatosensory cortex in adult flying-fox after digit amputation. *Nature*, 332(6163):446–448.

Callaway, E. M. et Katz, L. C. (1990). Emergence and refinement of clustered horizontal connections in cat striate cortex. *J. Neurosci.*, 10(4):1134–1153.

Callaway, E. M. et Katz, L. C. (1991). Effects of binocular deprivation on the development of clustered horizontal connections in cat striate cortex. *Proc. Natl. Acad. Sci. U.S.A.*, 88(3):745–749.

Capsoni, S., Tongiorgi, E., Cattaneo, A. et Domenici, L. (1999). Dark rearing blocks the developmental down-regulation of brain-derived neurotrophic factor messenger RNA expression in layers IV and V of the rat visual cortex. *Neuroscience*, 88(2):393–403.

Chapman, B. et Stryker, M. P. (1993). Development of orientation selectivity in ferret visual cortex and effects of deprivation. *J. Neurosci.*, 13(12):5251–5262.

Chapman, B., Stryker, M. P. et Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. *J. Neurosci.*, 16(20):6443–6453.

Chino, Y., Smith, 3rd, E. L., Zhang, B., Matsuura, K., Mori, T. et Kaas, J. H. (2001). Recovery of binocular responses by cortical neurons after early monocular lesions. *Nat. Neurosci.*, 4(7):689–690.

Chino, Y. M., Kaas, J. H., Smith, 3rd, E. L., Langston, A. L. et Cheng, H. (1992). Rapid reorganization of cortical maps in adult cats following restricted deafferentation in retina. *Vision Res.*, 32(5):789–796.

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J., Honda, M., Sadato, N., Gerloff, C., Catala, M. D. et Hallett, M. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389(6647):180–183.

Cohen, L. G., Weeks, R. A., Sadato, N., Celnik, P., Ishii, K. et Hallett, M. (1999). Period of susceptibility for cross-modal plasticity in the blind. *Ann. Neurol.*, 45(4):451–460.

Collignon, O., Renier, L., Bruyer, R., Tranduy, D. et Veraart, C. (2006). Improved selective and divided spatial attention in early blind subjects. *Brain Res.*, 1075(1):175–182.

Colonnese, M. T. et Constantine-Paton, M. (2001). Chronic NMDA receptor blockade from birth increases the sprouting capacity of ipsilateral retinocollicular axons without disrupting their early segregation. *J. Neurosci.*, 21(5):1557–1568.

Crair, M. C., Gillespie, D. C. et Stryker, M. P. (1998). The role of visual experience in the development of columns in cat visual cortex. *Science*, 279(5350):566–570.

Crair, M. C., Horton, J. C., Antonini, A. et Stryker, M. P. (2001). Emergence of ocular dominance columns in cat visual cortex by 2 weeks of age. *J. Comp. Neurol.*, 430(2):235–249.

Crair, M. C. et Malenka, R. C. (1995). A critical period for long-term potentiation at thalamocortical synapses. *Nature*, 375(6529):325–328.

Crowley, J. C. et Katz, L. C. (1999). Development of ocular dominance columns in the absence of retinal input. *Nat. Neurosci.*, 2(12):1125–1130.

Crowley, J. C. et Katz, L. C. (2000). Early development of ocular dominance columns. *Science*, 290(5495):1321–1324.

Cruikshank, S. J. et Weinberger, N. M. (1996). Evidence for the hebbian hypothesis in experience-dependent physiological plasticity of neocortex : a critical review. *Brain Res. Brain Res. Rev.*, 22(3):191–228.

Dammerman, R. S., Flint, A. C., Noctor, S. et Kriegstein, A. R. (2000). An excitatory gabaergic plexus in developing neocortical layer 1. *J. Neurophysiol.*, 84(1):428–434.

Diamond, D. M. et Weinberger, N. M. (1986). Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. *Brain Res.*, 372(2):357–360.

Doucet, M. E., Guillemot, J. P., Lassonde, M., Gagne, J. P., Leclerc, C. et Lepore, F. (2005). Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Exp. Brain Res.*, 160(2):194–202.

Elbert, T., Flor, H., Birbaumer, N., Knecht, S., Hampson, S., Larbig, W. et Taub, E. (1994). Extensive reorganization of the somatosensory cortex in adult humans after nervous system injury. *Neuroreport*, 5(18): 2593–2597.

Elbert, T., Heim, S. et Rockstroh, B. (2001). Neural plasticity and development. *In* Nelson, C. et Luciana, M., éditeurs : *Handbook of developmental cognitive neuroscience*, pages 191–202. Cambridge, MA : MIT Press.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B. et Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, 270(5234):305–307.

Epelbaum, M., Milleret, C., Buisseret, P. et Dufier, J. L. (1993). The sensitive period for strabismic amblyopia in humans. *Ophthalmology*, 100(3):323–327.

Ergenzinger, E. R., Glasier, M. M., Hahm, J. O. et Pons, T. P. (1998). Cortically induced thalamic plasticity in the primate somatosensory system. *Nat. Neurosci.*, 1(3):226–229.

Eysel, U. T., Schweigart, G., Mittmann, T., Eyding, D., Qu, Y., Vandesande, F., Orban, G. et Arckens, L. (1999). Reorganization in the visual cortex after retinal and cortical damage. *Restor. Neurol. Neurosci.*, 15(2-3):153–164.

Fagiolini, M. et Hensch, T. K. (2000). Inhibitory threshold for critical-period activation in primary visual cortex. *Nature*, 404(6774):183–186.

Fahle, M. et Poggio, T. (2002). *Perceptual learning*. Cambridge, MA : MIT Press.

Feldman, D. E. (2000). Inhibition and plasticity. *Nat. Neurosci.*, 3(4):303–304.

Ferster, D. et Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.*, 23:441–471.

Fine, I., Finney, E. M., Boynton, G. M. et Dobkins, K. R. (2005). Comparing the effects of auditory deprivation and sign language within the auditory and visual cortex. *J. Cogn. Neurosci.*, 17(10):1621–1637.

Finney, E. M., Clementz, B. A., Hickok, G. et Dobkins, K. R. (2003). Visual stimuli activate auditory cortex in deaf subjects : evidence from MEG. *Neuroreport*, 14(11):1425–1427.

Finney, E. M., Fine, I. et Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat. Neurosci.*, 4(12):1171–1173.

Flor, H., Elbert, T., Knecht, S., Wienbruch, C., Pantev, C., Birbaumer, N., Larbig, W. et Taub, E. (1995). Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. *Nature*, 375(6531):482–484.

Forster, C. R. et Illing, R. B. (1998). Redistribution of NMDA receptors in the cochlear nucleus following cochleotomy. *Neuroreport*, 9(15):3531–3535.

Fregnac, Y. et Imbert, M. (1984). Development of neuronal selectivity in primary visual cortex of cat. *Physiol. Rev.*, 64(1):325–434.

Frost, D. O. (1982). Anomalous visual connections to somatosensory and auditory systems following brain lesions in early life. *Brain Res.*, 255(4):627–635.

Frost, D. O., Boire, D., Gingras, G. et Ptito, M. (2000). Surgically created neural pathways mediate visual pattern discrimination. *Proc. Natl. Acad. Sci. U.S.A.*, 97(20):11068–11073.

Fuchs, J. L. et Salazar, E. (1998). Effects of whisker trimming on GABA(A) receptor binding in the barrel cortex of developing and adult rats. *J. Comp. Neurol.*, 395(2):209–216.

Galuske, R. A., Kim, D. S. et Singer, W. (1999). The role of neurotrophins in developmental cortical plasticity. *Restor. Neurol. Neurosci.*, 15(2-3):115–124.

Gao, W. J. et Pallas, S. L. (1999). Cross-modal reorganization of horizontal connectivity in auditory cortex without altering thalamocortical projections. *J. Neurosci.*, 19(18):7940–7950.

Garraghty, P. E., LaChica, E. A. et Kaas, J. H. (1991). Injury-induced reorganization of somatosensory cortex is accompanied by reductions in GABA staining. *Somatosens. Mot. Res.*, 8(4):347–354.

Gilbert, C. D. (1995). Dynamic properties of adult visual cortex. *In* Gazzaniga, M. S., éditeur : *The cognitive neurosciences*, pages 73–90. Cambridge, MA : MIT Press.

Gilbert, C. D., Sigman, M. et Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5):681–697.

Gilbert, C. D. et Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.*, 9(7):2432–2442.

Goldberg, M. C., Maurer, D., Lewis, T. L. et Brent, H. P. (2001). The influence of binocular visual deprivation on the development of visual-spatial attention. *Dev. Neuropsychol.*, 19(1):53–81.

Goldreich, D. et Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *J. Neurosci.*, 23(8):3439–3445.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J. et Belin, P. (2004). Neuropsychology : pitch discrimination in the early blind. *Nature*, 430(6997):309.

Grafman, J. (2000). Conceptualizing functional neuroplasticity. *J. Commun. Disord.*, 33(4):345–356.

Grant, A. C., Thiagarajah, M. C. et Sathian, K. (2000). Tactile perception in blind Braille readers : a psychophysical study of acuity and hyperacuity using gratings and dot patterns. *Percept. Psychophys.*, 62(2): 301–312.

Grusser, S. M., Winter, C., Muhlnickel, W., Denke, C., Karl, A., Villringer, K. et Flor, H. (2001). The relationship of perceptual phenomena and cortical reorganization in upper extremity amputees. *Neuroscience*, 102(2):263–272.

Gyllensten, L., Malmfors, T. et Norrlin, M. L. (1966). Growth alteration in the auditory cortex of visually deprived mice. *J. Comp. Neurol.*, 126(3):463–469.

Hamilton, R., Keenan, J. P., Catala, M. et Pascual-Leone, A. (2000). Alexia for Braille following bilateral occipital stroke in an early blind woman. *Neuroreport*, 11(2):237–240.

Hamilton, R. H., Pascual-Leone, A. et Schlaug, G. (2004). Absolute pitch in blind musicians. *Neuroreport*, 15(5):803–806.

Hardie, N. A., Martsi-McClintock, A., Aitkin, L. M. et Shepherd, R. K. (1998). Neonatal sensorineural hearing loss affects synaptic density in the auditory midbrain. *Neuroreport*, 9(9):2019–2022.

Harrison, R. V., Ibrahim, D., Stanton, S. G. et Mount, R. J. (1996). Reorganization of frequency maps in chinchilla auditory midbrain after longterm basal cochlear lesions induced at birth. *In* Salvi, R. J., Henderson, D., Fiorino, F. et Colletti, V., éditeurs : *Auditory System Plasticity and Regeneration*, pages 238–255. New York : Thieme Medical.

Harrison, R. V., Nagasawa, A., Smith, D. W., Stanton, S. et Mount, R. J. (1991). Reorganization of auditory cortex after neonatal high frequency cochlear hearing loss. *Hear. Res.*, 54(1):11–19.

Hashisaki, G. T. et Rubel, E. W. (1989). Effects of unilateral cochlea removal on anteroventral cochlear nucleus neurons in developing gerbils. *J. Comp. Neurol.*, 283(4):5–73.

Heinrichs, R. W. et Moorhouse, J. A. (1969). Touch-perception thresholds in blind diabetic subjects in relation to the reading of Braille type. *N. Engl. J. Med.*, 280(2):72–75.

Hendry, S. H. et Jones, E. G. (1986). Reduction in number of immunostained gabaergic neurones in deprived-eye dominance columns of monkey area 17. *Nature*, 320(6064):750–753.

Hickmott, P. W. et Constantine-Paton, M. (1997). Experimental down-regulation of the NMDA channel associated with synapse pruning. *J. Neurophysiol.*, 78(2):1096–1107.

Hoke, E. S., Hoke, M. et Ross, B. (1996). Neurophysiological correlate of the auditory after-image ("Zwicker tone"). *Audiol. Neurootol.*, 1(3):161–174.

Hoke, M., Feldmann, H., Pantev, C., Lütkenhöner, B. et Lehnertz, K. (1989). Objective evidence of tinnitus in auditory evoked magnetic fields. *Hear. Res.*, 37(3):281–286.

Hsieh, C. Y., Cruikshank, S. J. et Metherate, R. (2000). Differential modulation of auditory thalamocortical and intracortical synaptic transmission by cholinergic agonist. *Brain Res.*, 880(1-2):51–64.

Hubel, D. H. et Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.*, 195(1):215–243.

Hubel, D. H. et Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol.*, 206(2):419–436.

Irvine, D. R. (2000). Injury- and use-related plasticity in the adult auditory system. *J. Commun. Disord.*, 33(4):293–312.

Ito, M. (1986). Long-term depression as a memory process in the cerebellum. *Neurosci. Res.*, 3(6):531–539.

Jean-Baptiste, M. et Morest, D. K. (1975). Transynaptic changes of synaptic endings and nuclear chromitin in the trapezoid body following cochlear ablations in cats. *J. Comp. Neurol.*, 162:111–133.

Jenkins, W. M., Merzenich, M. M., Ochs, M. T., Allard, T. et Guic-Robles, E. (1990). Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J. Neurophysiol.*, 63(1):82–104.

Jones, E. G. et Pons, T. P. (1998). Thalamic and brainstem contributions to large-scale plasticity of primate somatosensory cortex. *Science*, 282(5391):1121–1125.

Kaas, J. H. (1991). Plasticity of sensory and motor maps in adult mammals. *Annu. Rev. Neurosci.*, 14:137–167.

Kaas, J. H. (1999). Is most of neural plasticity in the thalamus cortical? *Proc. Natl. Acad. Sci. U.S.A.*, 96(14):7622–7623.

Kaas, J. H. (2000). The reorganization of sensory and motor maps after injury in adult mammals. *In* Gazzaniga, M. S., éditeur : *The New Cognitive Neurosciences*, pages 223–236. Cambridge, MA : MIT Press.

Kaas, J. H. (2001). The mutability of sensory representations after injury in adult mammals. *In* Shaw, C. A. et McEachern, J. C., éditeurs : *Toward a theory of neuroplasticity*, pages 323–334. Hove, England : Psychology Press.

Kaas, J. H., Krubitzer, L. A., Chino, Y. M., Langston, A. L., Polley, E. H. et Blair, N. (1990). Reorganization of retinotopic cortical maps in adult mammals after lesions of the retina. *Science*, 248(4952):229–231.

Kaczmarek, L., Kossut, M. et Skangiel-Kramska, J. (1997). Glutamate receptors in cortical plasticity : molecular and cellular biology. *Physiol. Rev.*, 77(1):217–255.

Kaltenbach, J. A., Czaja, J. M. et Kaplan, C. R. (1992). Changes in the tonotopic map of the dorsal cochlear nucleus following induction of cochlear lesions by exposure to intense sound. *Hear. Res.*, 59(2):213–223.

Katz, L. C. et Shatz, C. J. (1996). Synaptic activity and the construction of cortical circuits. *Science*, 274(5290): 1133–1138.

Keech, R. V. et Kutschke, P. J. (1995). Upper age limit for the development of amblyopia. *J. Pediatr. Ophthalmol. Strabismus*, 32(2):89–93.

Kilgard, M. P. et Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science*, 279(5357):1714–1718.

Kim, D. S. et Bonhoeffer, T. (1994). Reverse occlusion leads to a precise restoration of orientation preference maps in visual cortex. *Nature*, 370(6488):370–372.

King, A. J. et Parsons, C. H. (1999). Improved auditory spatial acuity in visually deprived ferrets. *Eur. J. Neurosci.*, 11(11):3945–3956.

King, A. J. et Semple, D. J. (1996). Improvement in auditory spatial acuity following early visual deprivation in ferrets. *Soc. Neurosci. Abstr.*

Kiorpes, L., Kiper, D. C., O'Keefe, L. P., Cavanaugh, J. R. et Movshon, J. A. (1998). Neuronal correlates of amblyopia in the visual cortex of macaque monkeys with experimental strabismus and anisometropia. *J. Neurosci.*, 18(16):6411–6424.

Kirkwood, A. (2000). Serotonergic control of developmental plasticity. *Proc. Natl. Acad. Sci. U.S.A.*, 97(5):1951–1952.

Kirkwood, A., Rioult, M. C. et Bear, M. F. (1996). Experience-dependent modification of synaptic plasticity in visual cortex. *Nature*, 381(6582):526–528.

Kitzes, L. M. (1984). Some physiological consequences of neonatal cochlear destruction in the inferior colliculus of the gerbil, meriones unguiculatus. *Brain Res.*, 306(1-2):171–178.

Kitzes, L. M., Kageyama, G. H., Semple, M. N. et Kil, J. (1995). Development of ectopic projections from the ventral cochlear nucleus to the superior olivary complex induced by neonatal ablation of the contralateral cochlea. *J. Comp. Neurol.*, 353(3):341–363.

Kitzes, L. M. et Semple, M. N. (1985). Single-unit responses in the inferior colliculus : effects of neonatal unilateral cochlear ablation. *J. Neurophysiol.*, 53(6):1483–1500.

Knudsen, E. I. (1988). Sensitive and critical periods in the development of sound localization. *In* Easter, S. S., Barald, K. F. et Carlson, B. M., éditeurs : *From message to mind : Directions in developmental neurobiology*. Sunderland, MA : Sinauer Associates.

Kolb, B. et Gibb, R. (2001). Early brain injury, plasticity and behavior. *In* Nelson, C. et Luciana, M., éditeurs : *Handbook of developmental cognitive neuroscience*, pages 175–190. Cambridge, MA : MIT Press.

Kolb, B. et Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology*. New York : Freeman.

Kolb, B. et Whishaw, I. Q. (1998). Brain plasticity and behavior. *Annu. Rev. Psychol.*, 49:43–64.

Kral, A., Hartmann, R., Tillein, J., Heid, S. et Klinke, R. (2002). Hearing after congenital deafness : central auditory plasticity and sensory deprivation. *Cereb. Cortex*, 12(8):797–807.

Kral, A., Tillein, J., Heid, S., Klinke, R. et Hartmann, R. (2006). Cochlear implants : cortical plasticity in congenital deprivation. *Prog. Brain Res.*, 157:283–313.

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K. et Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *J. Cogn. Neurosci.*, 7:25.

Kujala, T., Huotilainen, M., Sinkkonen, J., Ahonen, A. I., Alho, K., Hamalainen, M. S., Ilmoniemi, R. J., Kajola, M., Knuutila, J. E., Lavikainen, J., Salonen, O., Simola, J., Standertskjöld-Nordenstam, C. G., Tiitinen, H., Tissari, S. O. et Näätänen, R. (1995). Visual cortex activation in blind humans during sound discrimination. *Neurosci. Lett.*, 183(1-2):143–146.

Kujala, T., Palva, M. J., Salonen, O., Alku, P., Huotilainen, M., Jarvinen, A. et Näätänen, R. (2005). The role of blind humans' visual cortex in auditory change detection. *Neurosci. Lett.*, 379(2):127–131.

Lambertz, N., Gizewski, E. R., de Greiff, A. et Forsting, M. (2005). Cross-modal plasticity in deaf subjects dependent on the extent of hearing loss. *Brain Res. Cogn. Brain Res.*, 25(3):884–890.

Le Grand, R., Mondloch, C. J., Maurer, D. et Brent, H. P. (2001). Neuroperception. early visual experience and face processing. *Nature*, 410(6831):890.

Lessard, N., Paré, M., Lepore, F. et Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395(6699):278–280.

Levänen, S. et Hamdorf, D. (2001). Feeling vibrations : enhanced tactile sensitivity in congenitally deaf humans. *Neurosci. Lett.*, 301(1):75–77.

Levänen, S., Jousmaki, V. et Hari, R. (1998). Vibration-induced auditory cortex activation in a congenitally deaf adult. *Curr. Biol.*, 8:869–872.

Loke, W. H. et Song, S. (1991). Central and peripheral visual processing in hearing and nonhearing individuals. *Bull. Psychon. Soc.*, 29:437.

Maurer, D. et Lewis, T. L. (2001). Visual acuity and spatial contrast sensitivity : Normal development and underlying mechanisms. *In* Nelson, C. et Luciana, M., éditeurs : *Handbook of developmental cognitive neuroscience*, pages 237–251. Cambridge, MA : MIT Press.

Maurer, D., Lewis, T. L. et Brent, H. (1989). The effects of deprivation on human visual development. studies of children treated with cataracts. *In* Morrison, F. J., Lord, C. et Keating, D. P., éditeurs : *Applied developmental psychology*, pages 139–227. San Diego : Academic Press.

McMullen, N. T., Goldberger, B., Suter, C. M. et Glaser, E. M. (1988). Neonatal deafening alters nonpyramidal dendrite orientation in auditory cortex : a computer microscope study in the rabbit. *J. Comp. Neurol.*, 267(1):92–106.

Melzer, P., Morgan, V. L., Pickens, D. R., Price, R. R., Wall, R. S. et Ebner, F. F. (2001). Cortical activation during Braille reading is influenced by early visual experience in subjects with severe visual disability : a correlational fMRI study. *Hum. Brain Mapp.*, 14(3):186–195.

Menning, H., Roberts, L. E. et Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport*, 11(4):817–822.

Merzenich, M. M. et Jenkins, W. M. (1993). Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience. *J. Hand Ther.*, 6(2):89–104.

Mills, C. (1985). Perception of visual temporal patterns by deaf and hearing adults. *Bull. Psychon. Soc.*, 23:483.

Moore, D. R. (1990). Auditory brainstem of the ferret : early cessation of developmental sensitivity of neurons in the cochlear nucleus to removal of the cochlea. *J. Comp. Neurol.*, 302(4):810–823.

Moore, D. R. (1992). Trophic influences of excitatory and inhibitory synapses on neurones in the auditory brain stem. *Neuroreport*, 3(3):269–272.

Moore, D. R., France, S. J., McAlpine, D., Mossop, J. E. et Versnel, H. (1997). Plasticity of inferior colliculus and auditory cortex following unilateral deafening in adult ferrets. *In* Syka, J., éditeur : *Acoustical signal processing in the central auditory system*, pages 489–499. New York : Plenum.

Moore, D. R., King, A. J., McAlpine, D., Martin, R. L. et Hutchings, M. E. (1993). Functional consequences of neonatal unilateral cochlear removal. *Prog. Brain Res.*, 97:127–133.

Moore, D. R. et Kitzes, L. M. (1985). Projections from the cochlear nucleus to the inferior colliculus in normal and neonatally cochlea-ablated gerbils. *J. Comp. Neurol.*, 240(2):180–195.

Moore, D. R. et Kitzes, L. M. (1986). Cochlear nucleus lesions in the adult gerbil : effects on neurone responses in the contralateral inferior colliculus. *Brain Res.*, 373(1-2):268–274.

Moore, D. R. et Kowalchuk, N. E. (1988). Auditory brainstem of the ferret : effects of unilateral cochlear lesions on cochlear nucleus volume and projections to the inferior colliculus. *J. Comp. Neurol.*, 272(4):503–515.

Moore, J. K., Niparko, J. K., Miller, M. R. et Linthicum, F. H. (1994). Effect of profound hearing loss on a central auditory nucleus. *Am. J. Otol.*, 15(5):588–595.

Morris, J. S., Friston, K. J. et Dolan, R. J. (1998). Experience-dependent modulation of tonotopic neural responses in human auditory cortex. *Proc. Biol. Sci.*, 265(1397):649–657.

Muchnik, C., Efrati, M., Nemeth, E., Malin, M. et Hildesheimer, M. (1991). Central auditory skills in blind and sighted subjects. *Scand. Audiol.*, 20(1):19–23.

Muhlnickel, W., Elbert, T., Taub, E. et Flor, H. (1998). Reorganization of auditory cortex in tinnitus. *Proc. Natl. Acad. Sci. U.S.A.*, 95(17):10340–10343.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P. et Turner, R. (1998). Cerebral organization for language in deaf and hearing subjects : biological constraints and effects of experience. *Proc. Natl. Acad. Sci. U.S.A.*, 95(3):922–929.

Neville, H. J. et Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task : an event-related potential and behavioral study. II. congenitally deaf adults. *Brain Res.*, 405(2):268–283.

Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. et Kubo, T. (1999). Sign language 'heard' in the auditory cortex. *Nature*, 397(6715):116.

Nordeen, K. W., Killackey, H. P. et Kitzes, L. M. (1983). Ascending auditory projections to the inferior colliculus in the adult gerbil, meriones unguiculatus. *J. Comp. Neurol.*, 214(2):131–143.

Pantev, C. et Lütkenhöner, B. (2000). Magentoencephalographic studies of functional organization and plasticity of the human auditory system. *Clin. Neurophysiol.*, 17:130–142.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E. et Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678):811–814.

Pantev, C., Wollbrink, A., Roberts, L. E., Engelien, A. et Lütkenhöner, B. (1999). Short-term plasticity of the human auditory cortex. *Brain Res.*, 842(1):192–199.

Parasnis, I. (1992). Allocation of attention, reading skills, and deafness. *Brain Lang.*, 43(4):583–596.

Parasnis, I. et Samar, V. J. (1985). Parafoveal attention in congenitally deaf and hearing young adults. *Brain Cogn.*, 4(3):313–327.

Parasnis, I., Samar, V. J. et Berent, G. P. (2003). Deaf adults without attention deficit hyperactivity disorder display reduced perceptual sensitivity and elevated impulsivity on the Test of Variables of Attention (T.O.V.A.). *J. Speech Lang. Hear. Res.*, 46(5):1166–1183.

Pascual-Leone, A. et Hamilton, R. (2001). The metamodal organization of the brain. *Prog. Brain Res.*, 134:427–445.

Pascual-Leone, A. et Torres, F. (1993). Plasticity of the sensorimotor cortex representation of the reading finger in Braille readers. *Brain*, 116:1.

Pasic, T. R., Moore, D. R. et Rubel, E. W. (1994). Effect of altered neuronal activity on cell size in the medial nucleus of the trapezoid body and ventral cochlear nucleus of the gerbil. *J. Comp. Neurol.*, 348(1):111–120.

Pasic, T. R. et Rubel, E. W. (1989). Rapid changes in cochlear nucleus cell size following blockade of auditory nerve electrical activity in gerbils. *J. Comp. Neurol.*, 283(4):474–480.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D. et Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages : implications for the neural basis of human language. *Proc. Natl. Acad. Sci. U.S.A.*, 97(25):13961–13966.

Pizzorusso, T., Berardi, N., Rossi, F. M., Viegi, A., Venstrom, K., Reichardt, L. F. et Maffei, L. (1999). Trka activation in the rat visual cortex by antirat trka igg prevents the effect of monocular deprivation. *Eur. J. Neurosci.*, 11(1):204–212.

Poizner, H. et Tallal, P. (1987). Temporal processing in deaf signers. *Brain Lang.*, 30(1):52–62.

Pons, T. P., Garraghty, P. E., Ommaya, A. K., Kaas, J. H., Taub, E. et Mishkin, M. (1991). Massive cortical reorganization after sensory deafferentation in adult macaques. *Science*, 252(5014):1857–1860.

Popelar, J., Erre, J. P., Aran, J. M. et Cazals, Y. (1994). Plastic changes in ipsi-contralateral differences of auditory cortex and inferior colliculus evoked potentials after injury to one ear in the adult guinea pig. *Hear. Res.*, 72(1-2):125–134.

Powell, T. P. et Erulkar, S. D. (1962). Transneuronal cell degeneration in the auditory relay nuclei of the cat. *J. Anat*, 96:249–268.

Proksch, J. et Bavelier, D. (2002). Changes in the spatial distribution of visual attention after early deafness. *J. Cogn. Neurosci.*, 14(5):687–701.

Quinlan, E. M., Olstein, D. H. et Bear, M. F. (1999a). Bidirectional, experience-dependent regulation of N-methyl-D-aspartate receptor subunit composition in the rat visual cortex during postnatal development. *Proc. Natl. Acad. Sci. U.S.A.*, 96(22):12876–12880.

Quinlan, E. M., Philpot, B. D., Huganir, R. L. et Bear, M. F. (1999b). Rapid, experience-dependent expression of synaptic NMDA receptors in visual cortex in vivo. *Nat. Neurosci.*, 2(4):352–357.

Rajan, R. et Irvine, D. R. (1998). Absence of plasticity of the frequency map in dorsal cochlear nucleus of adult cats after unilateral partial cochlear lesions. *J. Comp. Neurol.*, 399(1):35–46.

Rajan, R., Irvine, D. R., Wise, L. Z. et Heil, P. (1993). Effect of unilateral partial cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J. Comp. Neurol.*, 338(1):17–49.

Rauschecker, J. P. et Kniepert, U. (1994). Auditory localization behaviour in visually deprived cats. *Eur. J. Neurosci.*, 6(1):149–160.

Rauschecker, J. P., Tian, B., Korte, M. et Egert, U. (1992). Crossmodal changes in the somatosensory vibrissa/barrel system of visually deprived animals. *Proc. Natl. Acad. Sci. U.S.A.*, 89(11):5063–5067.

Röder, B. et Neville, H. (2003). Developmental functional plasticity. *In* Grafman, J. et Robertson, I., éditeurs : *Handbook of Neurospychology*. New York : Elsevier.

Reale, R. A., Brugge, J. F. et Chan, J. C. (1987). Maps of auditory cortex in cats reared after unilateral cochlear ablation in the neonatal period. *Brain Res.*, 431(2):281–290.

Recanzone, G. H., Jenkins, W. M., Hradek, G. T. et Merzenich, M. M. (1992a). Progressive improvement in discriminative abilities in adult owl monkeys performing a tactile frequency discrimination task. *J. Neurophysiol.*, 67(5):1015–1030.

Recanzone, G. H., Merzenich, M. M. et Jenkins, W. M. (1992b). Frequency discrimination training engaging a restricted skin surface results in an emergence of a cutaneous response zone in cortical area 3a. *J. Neurophysiol.*, 67(5):1057–1070.

Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, K. A. et Dinse, H. R. (1992c). Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *J. Neurophysiol.*, 67(5):1031–1056.

Recanzone, G. H., Merzenich, M. M. et Schreiner, C. E. (1992d). Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task. *J. Neurophysiol.*, 67(5):1071–1091.

Recanzone, G. H., Schreiner, C. E. et Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.*, 13(1):87–103.

Regal, D. M., Boothe, R., Teller, D. Y. et Sackett, G. P. (1976). Visual acuity and visual responsiveness in dark-reared monkeys (macaca nemestrina). *Vision Res.*, 16(5):523–530.

Rettenbach, R., Diller, G. et Sireteanu, R. (1999). Do deaf people see better ? texture segmentation and visual search compensate in adult but not in juvenile subjects. *J. Cogn. Neurosci.*, 11(5):560–583.

Rice, C. E. (1970). Early blindness, early experience and perceptual enhancement. *Res. Bull. Am. Found. Blind*, 22:1–22.

Rice, C. E., Feinstein, S. H. et Schusterman, R. J. (1965). Echo-detection ability of the blind : size and distance factors. *J. Exp. Psychol.*, 70:246–255.

Rittenhouse, C. D., Shouval, H. Z., Paradiso, M. A. et Bear, M. F. (1999). Monocular deprivation induces homosynaptic long-term depression in visual cortex. *Nature*, 397(6717):347–350.

Robertson, D. et Irvine, D. R. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.*, 282(3):456–471.

Robertson, R. T. (1987). A morphogenic role for transiently expressed acetylcholinesterase in developing thalamocortical systems ? *Neurosci. Lett.*, 75(3):259–264.

Robertson, R. T., Mostamand, F., Kageyama, G. H., Gallardo, K. A. et Yu, J. (1991). Primary auditory cortex in the rat : transient expression of acetylcholinesterase activity in developing geniculocortical projections. *Brain Res. Dev. Brain Res.*, 58(1):81–95.

Roder, B. et Rosler, F. (2004). Compensatory plasticity as a consequence of sensory loss. *In* Calvert, G. A., Spence, C. et Stein, B. E., éditeurs : *The Handbook of Multisensory Processes*, pages 719–747. Cambridge, MA : MIT Press.

Roder, B., Stock, O., Bien, S., Neville, H. et Rosler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *Eur. J. Neurosci.*, 16(5):930–936.

Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S. A. et Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, 400(6740):162–166.

Roe, A. W., Garraghty, P. E., Esguerra, M. et Sur, M. (1993). Experimentally induced visual projections to the auditory thalamus in ferrets : evidence for a W cell pathway. *J. Comp. Neurol.*, 334(2):263–280.

Roe, A. W., Hahm, J.-O. et Sur, M. (1991). Experimentally induced establishment of visual topography in auditory thalamus. *Soc. Neurosci. Abstr.*, 17:898.

Roe, A. W., Pallas, S. L., Hahm, J. O. et Sur, M. (1990). A map of visual space induced in primary auditory cortex. *Science*, 250(4982):818–820.

Roe, A. W., Pallas, S. L., Kwon, Y. H. et Sur, M. (1992). Visual projections routed to the auditory pathway in ferrets : receptive fields of visual neurons in primary auditory cortex. *J. Neurosci.*, 12(9):3651–3664.

Russell, F. A. et Moore, D. R. (1995). Afferent reorganisation within the superior olivary complex of the gerbil : development and induction by neonatal, unilateral cochlear removal. *J. Comp. Neurol.*, 352(4):607–625.

Ruthazer, E. S. et Stryker, M. P. (1996). The role of activity in the development of long-range horizontal connections in area 17 of the ferret. *J. Neurosci.*, 16(22):7253–7269.

Ryugo, D. K., Ryugo, R., Globus, A. et Killackey, H. P. (1975). Increased spine density in auditory cortex following visual or somatic deafferentation. *Brain Res.*, 90(1):143–146.

Sadato, N., Okada, T., Honda, M. et Yonekura, Y. (2002). Critical period for cross-modal plasticity in blind humans : a functional MRI study. *Neuroimage*, 16(2):389–400.

Sadato, N., Pascual-Leone, A., Grafman, J., Deiber, M. P., Ibanez, V. et Hallett, M. (1998). Neural networks for Braille reading by the blind. *Brain*, 121:7.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M. P., Dold, G. et Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380(6574):526–528.

Sadato, N., Yamada, H., Okada, T., Yoshida, M., Hasegawa, T., Matsuki, K., Yonekura, Y. et Itoh, H. (2004). Age-dependent plasticity in the superior temporal sulcus in deaf humans : a functional MRI study. *BMC Neurosci.*, 5(1):56.

Sathian, K. (1998). Perceptual learning. *Curr. Sci.*, 75:451–457.

Schneider, G. E. (1973). Early lesions of superior colliculus : factors affecting the formation of abnormal retinal projections. *Brain Behav. Evol.*, 8(1):73–109.

Schroeder, C. E., Smiley, J., Fu, K. G., McGinnis, T., O'Connell, M. N. et Hackett, T. A. (2003). Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *Int. J. Psychophysiol.*, 50(1-2):5–17.

Schwaber, M. K., Garraghty, P. E. et Kaas, J. H. (1993). Neuroplasticity of the adult primate auditory cortex following cochlear hearing loss. *Am. J. Otol.*, 14(3):252–258.

Segal, I., Korkotian, I. et Murphy, D. D. (2000). Dendritic spine formation and pruning : common cellular mechanisms ? *Trends Neurosci.*, 23(2):53–57.

Sermasi, E., Tropea, D. et Domenici, L. (1999). A new form of synaptic plasticity is transiently expressed in the developing rat visual cortex : a modulatory role for visual experience and brain-derived neurotrophic factor. *Neuroscience*, 91(1):163–173.

Sharma, J., Angelucci, A. et Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, 404(6780):841–847.

Shatz, C. J. et Stryker, M. P. (1978). Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *J. Physiol.*, 281:267–283.

Sherman, S. M. et Spear, P. D. (1982). Organization of visual pathways in normal and visually deprived cats. *Physiol. Rev.*, 62(2):738–855.

Sladen, D. P., Tharpe, A. M., Ashmead, D. H., Wesley Grantham, D. et Chun, M. M. (2005). Visual attention in deaf and normal hearing adults : effects of stimulus compatibility. *J. Speech Lang. Hear. Res.*, 48(6):1529–1537.

Somers, D. C., Dragoi, V. et Sur, M. (2001). Orientation-selectivity and its modulation by local and long-range connections in visual cortex. *In* Peters, A. et Payne, B., éditeurs : *Cerebral cortex : The cat primary visual cortex.*, pages 471–520. San Diego, CA : Academic Press.

Somers, D. C., Nelson, S. B. et Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.*, 15(8):5448–5465.

Stanton, S. G. et Harrison, R. V. (2000). Projections from the medial geniculate body to primary auditory cortex in neonatally deafened cats. *J. Comp. Neurol.*, 426(1):117–129.

Sterr, A., Muller, M. M., Elbert, T., Rockstroh, B., Pantev, C. et Taub, E. (1998). Perceptual correlates of changes in cortical representation of fingers in blind multifinger Braille readers. *J. Neurosci.*, 18(11):4417–4423.

Stevens, A. A. et Weaver, K. (2005). Auditory perceptual consolidation in early-onset blindness. *Neuropsychologia*, 43(13):1901–1910.

Stevens, C. et Neville, H. (2006). Neuroplasticity as a double-edged sword : deaf enhancements and dyslexic deficits in motion processing. *J. Cogn. Neurosci.*, 18(5):701–714.

Stevens, J. C., Foulke, E. et Patterson, M. Q. (1996). Tactile acuity, aging, and Braille reading in long-term blindness. *J. Exp. Psychol. Appl.*, 2:91.

Stivalet, P., Moreno, Y., Richard, J., Barraud, P. A. et Raphel, C. (1998). Differences in visual search tasks between congenitally deaf and normally hearing adults. *Brain Res. Cogn. Brain Res.*, 6(3):227–232.

Sur, M. (2004). Rewiring cortex : cross-modal plasticity and its implications for cortical development and function. *In* Calvert, G., Spence, C. et Stein, B. E., éditeurs : *The handbook of multisensory processes*, pages 681–694. Cambridge, MA : MIT Press.

Sur, M., Angelucci, A. et Sharma, J. (1999). Rewiring cortex : the role of patterned activity in development and plasticity of neocortical circuits. *J. Neurobiol.*, 41(1):33–43.

Sur, M., Garraghty, P. E. et Roe, A. W. (1988). Experimentally induced visual projections into auditory thalamus and cortex. *Science*, 242(4884):1437–1441.

Sur, M. et Leamey, C. A. (2001). Development and plasticity of cortical areas and networks. *Nat. Rev. Neurosci.*, 2(4):251–262.

Sur, M., Pallas, S. L. et Roe, A. W. (1990). Cross-modal plasticity in cortical development : differentiation and specification of sensory neocortex. *Trends Neurosci.*, 13(6):227–233.

Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiol. Rev.*, 82(3):601–636.

Tang, Y. P., Shimizu, E., Dube, G. R., Rampon, C., Kerchner, G. A., Zhuo, M., Liu, G. et Tsien, J. Z. (1999). Genetic enhancement of learning and memory in mice. *Nature*, 401(6748):63–69.

Teyler, T. J. (2001). LTP and the superfamily of synaptic plasticities. *In* Shaw, C. A. et McEachern, J. C., éditeurs : *Toward a theory of neuroplasticity*, pages 101–117. Hove, England : Psychology Press.

Tierney, T. S., Russell, F. A. et Moore, D. R. (1997). Abrupt cessation during early postnatal development of afferent dependent survival of gerbil cochlear nucleus neurons. *J. Comp. Neurol.*, 378:295.

Toldi, J., Farkas, T. et Volgyi, B. (1994). Neonatal enucleation induces cross-modal changes in the barrel cortex of rat. A behavioural and electrophysiological study. *Neurosci. Lett.*, 167(1-2):1–4.

Turner, A. M. et Greenough, W. T. (1985). Differential rearing effects on rat visual cortex synapses. I. synaptic and neuronal density and synapses per neuron. *Brain Res.*, 329(1-2):195–203.

Uhl, F., Franzen, P., Lindinger, G., Lang, W. et Deecke, L. (1991). On the functionality of the visually deprived occipital cortex in early blind persons. *Neurosci. Lett.*, 124(2):256–259.

Vaegan, T. D. et Taylor, D. (1979). Critical period for deprivation amblyopia in children. *Trans. Ophthalmol. Soc. U.K.*, 99(3):432–439.

Van Boven, R. W., Hamilton, R. H., Kauffman, T., Keenan, J. P. et Pascual-Leone, A. (2000). Tactile spatial resolution in blind Braille readers. *Neurology*, 54(12):2230–2236.

von Melchner, L., Pallas, S. L. et Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, 404(6780):871–876.

Von Noorden, G. K. (1972). Experimental amblyopia. *Isr. J. Med. Sci.*, 8(8):1496–1499.

Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J. P. et Lepore, F. (2004). Early- and late-onset blind individuals show supra-normal auditory abilities in far-space. *Curr. Biol.*, 14(19):1734–1738.

Wanet-Defalque, M. C., Veraart, C., De Volder, A., Metz, R., Michel, C., Dooms, G. et Goffinet, A. (1988). High metabolic activity in the visual cortex of early blind human subjects. *Brain Res.*, 446(2):369–373.

Wang, X., Merzenich, M. M., Sameshima, K. et Jenkins, W. M. (1995). Remodelling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature*, 378(6552):71–75.

Warren, R., Tremblay, N. et Dykes, R. W. (1989). Quantitative study of glutamic acid decarboxylase-immunoreactive neurons and cytochrome oxidase activity in normal and partially deafferented rat hindlimb somatosensory cortex. *J. Comp. Neurol.*, 288(4):583–592.

Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., Hallett, M. et Rauschecker, J. P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *J. Neurosci.*, 20(7):2664–2672.

Weinberger, N. M. (1995). Dynamic regulation of receptive fields and maps in the adult sensory cortex. *Annu. Rev. Neurosci.*, 18:129–158.

Weinberger, N. M., Javid, R. et Lepan, B. (1993). Long-term retention of learning-induced receptive-field plasticity in the auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.*, 90(6):2394–2398.

Welker, E., Soriano, E. et Van der Loos, H. (1989). Plasticity in the barrel cortex of the adult mouse : effects of peripheral deprivation on GAD-immunoreactivity. *Exp. Brain Res.*, 74(3):441–452.

Wiesel, T. N. et Hubel, D. H. (1965). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *J. Neurophysiol.*, 28(6):1029–1040.

Wiesel, T. N. et Hubel, D. H. (1974). Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.*, 158(3):307–318.

Witkin, H. A., Birnbaum, J., Lomonaco, S., Lehr, S. et Herman, J. L. (1968). Cognitive patterning in congenitally totally blind children. *Child Dev.*, 39(3):766–786.

Xerri, C., Stern, J. M. et Merzenich, M. M. (1994). Alterations of the cortical representation of the rat ventrum induced by nursing behavior. *J. Neurosci.*, 14(3 Pt 2):1710–1721.

Yabe, T. et Kaga, K. (2005). Sound lateralization test in adolescent blind individuals. *Neuroreport*, 16(9):939–942.

Chapitre 9

Réorganisation neurofonctionnelle après implantation cochléaire

Sommaire

9.1 Effets de l'implantation sur la perception de la parole 192
9.1.1 Perception auditive de la parole après implantation 192
9.1.2 Perception visuelle de la parole après implantation 193
9.1.3 Perception audiovisuelle de la parole après implantation 193
9.2 Récupération perceptive : facteurs techniques et chirurgicaux . . . 194
9.2.1 Nombre d'électrodes et stratégie de codage 194
9.2.2 Rôle de l'enveloppe temporelle et de la structure fine 194
9.2.3 Influence de la profondeur d'implantation 196
9.3 Récupération perceptive : facteurs neurofonctionnels 198
9.3.1 Facteurs neurophysiologiques 198
9.3.2 Effet de l'expérience sensorielle 199
9.3.3 Période sensible pour l'implantation cochléaire 199
9.3.4 Utilisation du métabolisme pré-implantation comme facteur prédictif . . 200
9.4 Réorganisation corticale après implantation cochléaire 201
9.4.1 Réactivation des voies auditives centrales 201
9.4.2 Changements de spécialisation fonctionnelle 202
9.4.3 Sur- et sous-activations corticales 203
9.4.4 Interactions multisensorielles 203
Bibliographie 205

Introduction

La surdité et l'implantation cochléaire consistent chacune en une altération des entrées auditives parvenant aux centres nerveux. Dans le cas de la surdité, il s'agit d'une déprivation sensorielle, qui anéantit ou diminue considérablement la quantité d'informations sensorielles délivrée aux centres nerveux par le nerf auditif. À l'inverse, après une implantation cochléaire, le cortex reçoit un flot inhabituel d'informations en provenance du nerf auditif.

L'implant cochléaire ne restitue que très partiellement la tonotopie physiologique normale. Par ailleurs, la stimulation électrique appliquée au nerf auditif est encore assez éloignée de la stimulation normale produite par une cochlée saine (pour plus de détails sur le codage électrique des implants cochléaires, se référer à *« Aspects techniques »*, sect. 7.5, p. 149). Le message nerveux généré par l'implant cochléaire est donc considérablement dégradé et appauvri en information, en comparaison avec un système auditif intact.

Dans les deux cas, le cortex va devoir s'adapter à ces nouvelles afférences sensorielles, par l'intermédiaire des phénomènes naturels d'apprentissage et de réorganisation corticale. De par le déséquilibre survenu entre les afflux nerveux en provenance des différentes modalités sensorielles, de nouvelles associations fonctionnelles et neuronales vont se créer progressivement, aboutissant à des recrutements des aires cérébrales différents d'une répartition normale. C'est cette différence que nous nous proposons de détailler dans ce chapitre, qui sert de préambule à l'exposé de nos travaux correspondants (*« Etude des conséquences de l'implantation cochléaire »*, part. III, p. 217).

9.1 Effets de l'implantation sur la perception de la parole

Nous décrivons dans la suite les observations rapportées dans la littérature scientifique concernant l'évolution de la compréhension de la parole auditive, visuelle et audiovisuelle à la suite d'une implantation cochléaire. On pourra se référer à *« Effets de la surdité sur la perception de la parole »*, sect. 6.4, p. 137 pour une description de cette évolution dans le cas de la surdité.

9.1.1 Perception auditive de la parole après implantation

Les études portant sur la perception auditive de la parole chez les personnes sourdes implantées cochléaires s'accordent pour indiquer une nette amélioration de la reconnaissance de la parole auditive après implantation cochléaire. Ce fait est appuyé par des données expérimentales chez l'enfant sourd prélingual implanté (Bergeson *et al.*, 2003; Calmels *et al.*, 2004) et l'adulte sourd postlingual implanté (Chouard *et al.*, 1983; Wilson *et al.*, 1991; UK Cochlear Implant Study Group, 2004). Par ailleurs les performances auditives continuent à s'améliorer progressivement en fonction de la durée d'utilisation de l'implant cochléaire, aussi bien chez les enfants sourds prélinguaux (Bergeson *et al.*, 2003) que chez les adultes sourds postlinguaux (Pelizzone *et al.*, 1999; Hamzavi *et al.*, 2003).

De façon générale, l'analyse détaillée de la perception auditive avec un implant cochléaire permet de mettre en évidence une détérioration notable des indices auditifs de lieu d'articulation, ainsi qu'une dégradation des indices de voisement et de nasalité. Ces dégradations perceptives sont particulièrement bien illustrées par les nombreuses études portant sur les confusions auditives de consonnes chez les adultes sourds postlinguaux implantés cochléaires (Dowell *et al.*, 1982; Danhauer *et al.*, 1986; Tyler *et al.*, 1986; Blamey *et al.*, 1987; Rosen *et al.*, 1989; Dorman *et al.*, 1990; Rabinowitz *et al.*, 1992; McKay et McDermott, 1993; Doyle *et al.*, 1995; Whitford *et al.*, 1995; Helms *et al.*, 1997; Parkinson *et al.*, 1998; Pelizzone *et al.*, 1999; Kompis *et al.*, 1999; Skinner *et al.*, 1999; van Wieringen et Wouters, 1999; Ziese *et al.*, 2000; Kiefer *et al.*, 2001; Valimaa *et al.*, 2002).

Pour mémoire, concernant le sujet normo-entendant, on pourra se référer à *« Substrats neuronaux de la parole auditive »*, sect. 4.2, p. 81 pour une description des mécanismes neuronaux impliqués dans la perception et la compréhension de la parole auditive. On pourra également se référer à *« Perception auditive de la parole lors de la surdité »*, § 6.4.1, p. 137 pour une description des conséquences de la surdité sur la perception de la parole auditive.

9.1.2 Perception visuelle de la parole après implantation

Plusieurs études semblent montrer une légère amélioration de la reconnaissance visuelle de la parole (lecture labiale) à la suite d'une implantation cochléaire, à la fois au niveau du phonème et du mot entier. Ainsi Tyler *et al.* (1997) ont démontré une corrélation positive significative entre la durée d'utilisation de l'implant et les performances visuelles phonémiques chez des enfants sourds prélinguaux implantés cochléaires. Ces résultats ont été répliqués dans une population similaire avec des présentations visuelles de mots et de phrases (Bergeson *et al.*, 2003, 2005), ainsi que dans une population d'adultes sourds postlinguaux implantés cochléaires (Giraud *et al.*, 2001b). Toutefois cette amélioration longitudinale des performances visuelles se produit de façon plus lente par rapport aux améliorations auditives et audiovisuelles (Bergeson *et al.*, 2005).

Pour mémoire, concernant le sujet normo-entendant, on pourra se référer à *« Substrats neuronaux de la parole visuelle »*, sect. 4.3, p. 82 pour une description des mécanismes neuronaux impliqués dans la perception et la compréhension de la parole visuelle. On pourra également se référer à *« Perception visuelle de la parole lors de la surdité »*, § 6.4.2, p. 137 pour une description des conséquences de la surdité sur la perception de la parole visuelle.

9.1.3 Perception audiovisuelle de la parole après implantation

De façon générale, comme dans le cas des personnes normo-entendantes et déficientes auditives, la présentation simultanée des composantes auditive et visuelle de la parole aux utilisateurs d'implants cochléaires provoque une amélioration perceptive en comparaison avec une présentation unimodale.

Plusieurs études comparant les scores perceptifs en condition auditive, visuelle et audiovisuelle montrent ainsi une amélioration bimodale de l'identification des mots ou des phonèmes chez les enfants sourds prélinguaux ayant subi une implantation cochléaire (Staller *et al.*, 1991; Tyler et Lowder, 1992; Geers, 1994; Geers et Brenner, 1994; Tyler *et al.*, 1997; Blamey *et al.*, 2001; Lachs *et al.*, 2001; Geers *et al.*, 2003; Bergeson *et al.*, 2003, 2005). Ces données ont également été répliquées chez des adultes sourds postlinguaux implantés cochléaires (Goh *et al.*, 2001; Kaiser *et al.*, 2003), ainsi que dans le cas des adultes sourds prélinguaux implantés tardivement (Moody-Antonio *et al.*, 2005).

L'étude de Goh *et al.* (2001) est particulièrement intéressante puisque les expérimentateurs ont testé un patient sourd postlingual implanté cochléaire obtenant de très bons résultats comportementaux, le test se faisant en condition de parole sinusoïdale (similaire au vocoder, mais les enveloppes de bande sont modulées par des sons purs au lieu de bruit blanc). Alors que les contrôles normo-entendants ont montré de meilleures performances auditives, le patient implanté a obtenu de meilleurs résultats non seulement en condition visuelle seule mais également en condition audiovisuelle, indiquant par là une meilleure capacité d'intégration des informations auditives avec les informations visuelles.

De façon générale, comme dans le cas de la perception de parole auditive ou visuelle, on peut observer après implantation cochléaire une amélioration progressive des performances audiovisuelles en fonction de la durée d'utilisation de l'implant, comme le montrent les études longitudinales menées chez les enfants sourds prélinguaux implantés cochléaires (Bergeson *et al.*, 2003).

Pour mémoire, concernant le sujet normo-entendant, on pourra se référer à *« Perception audiovisuelle de la parole »*, ch. 3, p. 59 et à *« Réseau cérébral audio-visuo-moteur de la parole »*, sect. 4.4, p. 84 pour une description des mécanismes fonctionnels et neuronaux impliqués dans la perception et la compréhension de la parole audiovisuelle. On pourra également se référer à

« *Perception audiovisuelle de la parole lors de la surdité* », § 6.4.3, p. 138 pour une description des conséquences de la surdité sur la perception de la parole audiovisuelle.

9.2 Récupération perceptive : facteurs techniques et chirurgicaux

9.2.1 Nombre d'électrodes et stratégie de codage

De nombreuses études portant sur les différences entre les performances de compréhension de la parole après implantation cochléaire ont porté sur les paramètres liés au traitement du signal et au positionnement cochléaire des implants. Ces études ont permis de mettre en évidence un lien entre ces paramètres et certains aspects de la perception de la parole chez les patients implantés cochléaires (Blamey *et al.*, 1996a; Valimaa, 2002).

Plusieurs études ont ainsi montré que les patients implantés ont besoin d'un nombre plus grand d'électrodes stimulatrices pour la perception des voyelles (7 canaux) que pour la perception des consonnes ou des phrases (4 canaux) (Skinner *et al.*, 1995; Fishman *et al.*, 1997; Kiefer *et al.*, 2000). Inversement, une augmentation de la fréquence de stimulation électrique produit une plus grande amélioration de la compréhension pour les consonnes que pour les voyelles (Fu et Shannon, 2000; Kiefer *et al.*, 2000; Loizou *et al.*, 2000; Kiefer *et al.*, 2001). L'ampleur de la dynamique acoustique des différents canaux fréquentiels est nettement associée avec la qualité de la compréhension de la parole dans le bruit (Fu et Shannon, 1999a; Zeng et Galvin, 1999). Des études ont également montré une association entre les performances dans la compréhension auditive de la parole et la configuration, l'étendue fréquentielle globale et la répartition fréquentielle des électrodes (Fu et Shannon, 1999b; Friesen *et al.*, 1999), leur localisation et leur espacement dans la cochlée (Fu et Shannon, 1999c), ainsi que les réglages de l'amplification du signal par contrôle automatique du gain (Stobich *et al.*, 1999).

De nettes différences ont également été démontrées en fonction des stratégies de traitement du signal utilisées par le processeur de l'implant, aussi bien au niveau de la perception phonétique et de la compréhension du langage qu'au niveau de la perception de la hauteur des sons (Tyler *et al.*, 1997; Valimaa, 2002; Skinner *et al.*, 2002; David *et al.*, 2003; Beynon *et al.*, 2003; Loizou *et al.*, 2003; Beynon et Snik, 2004; Zeng, 2004; Manrique *et al.*, 2005; Vandali *et al.*, 2005; Zwolan *et al.*, 2005).

9.2.2 Rôle de l'enveloppe temporelle et de la structure fine

Le signal délivré en sortie d'un filtre auditif (voir « *Filtres auditifs et masquage psychoacoustique* », § 1.2.3, p. 27) peut être considéré comme le résultat d'un filtrage passe-bande du son original. Ce signal contient deux types d'information : d'une part les fluctuations temporelles globales de l'énergie acoustique, correspondant à l'enveloppe temporelle du signal, et d'autre part les oscillations rapides se produisant à une fréquence proche de celle du filtre auditif, ces dernières correspondant à la structure temporelle fine du signal (Rosen, 1992). Le signal acoustique issu d'un filtre auditif peut donc être vu comme une onde porteuse oscillant rapidement (la structure fine) modulée par une onde aux fluctuations relativement lentes (l'enveloppe temporelle). Les études du système auditif tendent à indiquer que l'enveloppe temporelle est codée par des fluctuations dans le taux de décharge des neurones, tandis que la structure fine serait codée par une synchronisation des décharges par rapport à une phase spécifique de l'onde porteuse (*phase locking*) (Rose *et al.*, 1967; Joris et Yin, 1992).

L'étude détaillée des rôles joués par l'enveloppe temporelle et la structure temporelle fine a pu être réalisée par une subdivision des sons en différentes bandes fréquentielles contigües, suivie

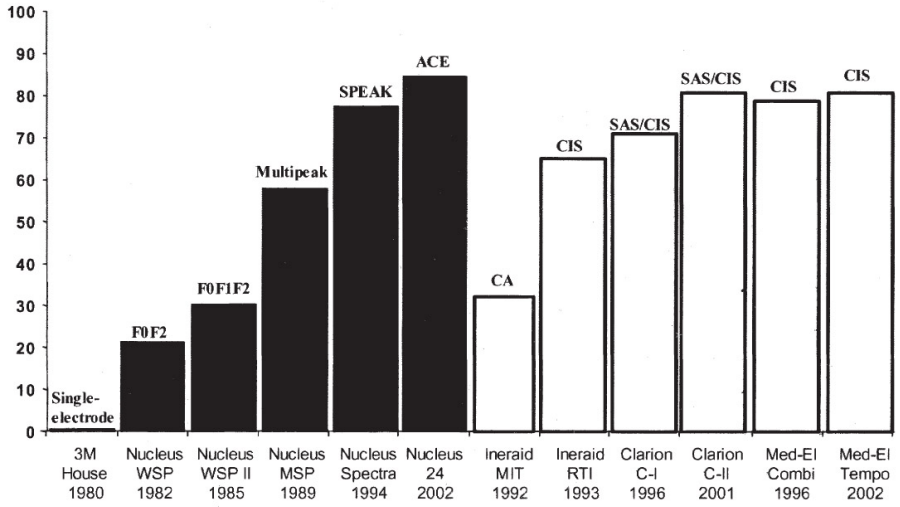


FIG. 9.1 – Compréhension de la parole auditive chez les patients implantés cochléaires en fonction des stratégies de traitement du son. L'échelle des ordonnées représente le pourcentage de phrases reconnues dans le silence. L'échelle des abscisses indique le constructeur, le type de processeur et l'année de publication de l'étude. Le texte au-dessus des histogrammes indique le type de stratégie de codage utilisée ; les stratégies spectrales et hybrides sont figurées en noir tandis que les stratégies temporelles sont figurées en blanc. Adapté d'après Zeng (2004).

par un traitement spécifique permettant de ne conserver que les indices associées à l'une ou l'autre. Ainsi la préservation des indices d'enveloppe temporelle peut être assurée par extraction de l'enveloppe du signal de bande, puis par utilisation de cette enveloppe pour moduler un bruit blanc de bande ou une son pur correspondant à la fréquence centrale de la bande d'origine. C'est le principe de base utilisé à l'origine par le *vocoder* (Dudley, 1939a,b), puis par les simulations acoustiques d'implant cochléaire (Shannon *et al.*, 1995). Les opérations de traitement du signal employées dans ces simulations sont décrites de façon détaillée dans « *Simulation acoustique de l'implant cochléaire* », sect. 10.2.2, p. 218. Les données expérimentales démontrent qu'un nombre de bandes spectrales modéré (entre 4 et 16) pour lesquelles seules les informations d'enveloppe temporelle sont préservées peut aboutir à des niveaux élevés de reconnaissance auditive de la parole chez les sujets normo-entendants (Shannon *et al.*, 1995; Loizou *et al.*, 1999) aussi bien que chez les personnes déficientes auditives (Turner *et al.*, 1995; Souza et Boike, 2006).

Ces observations indiquent que les informations d'enveloppe temporelle sont suffisantes pour assurer une compréhension convenable de la parole dans des conditions non bruitées. Toutefois l'ajout d'un bruit d'arrière-plan à des stimuli de parole ne conservant que les informations d'enveloppe (vocoder) provoque une forte dégradation des performances auditives chez les sujets normo-entendants (Nelson *et al.*, 2003; Qin et Oxenham, 2003; Stone et Moore, 2003; Zeng *et al.*, 2005; Fullgrabe *et al.*, 2006). On peut faire des observations identiques dans le cas des sujets implantés cochléaires, pour lesquels la compréhension de la parole se dégrade très rapidement avec le rapport signal/bruit (Fu *et al.*, 1998; Friesen *et al.*, 2001; Munson et Nelson, 2005). Cela suggère la grande importance du rôle joué par les indices de structure temporelle fine dans la compréhension de la parole en condition bruitée.

La préservation des indices de structure temporelle fine peut être réalisée en extrayant l'enveloppe du signal de bande grâce à une transfomée de Hilbert puis en divisant le signal de bande par son enveloppe temporelle (Smith *et al.*, 2002), ce qui aboutit à un signal dont l'enve-

loppe temporelle est constante et dont la structure fine est identique à celle du signal de bande original. Cette procédure nécessite l'utilisation d'un nombre de bandes spectrales suffisantes (16 ou plus) pour limiter la présence d'indices d'enveloppe résiduels (Lorenzi *et al.*, 2006), la corrélation entre indices d'enveloppe et indices de structure fine étant en effet d'autant plus marquée que le nombre de bandes spectrales est faible (Zeng *et al.*, 2004). L'utilisation de signaux dont seule la structure temporelle fine est préservée a ainsi permis de démontrer que les patients déficients auditifs, à l'inverse des sujets normo-entendants, ne sont pas capables d'utiliser les indices de structure fine pour comprendre la parole (Lorenzi *et al.*, 2006). Il est très probable que les patients implantés cochléaires se trouvent dans une situation identique, le traitement du signal acoustique effectué par l'implant ne préservant que les informations d'enveloppe temporelle, au détriment des informations de structure fine. Cela pourrait expliquer la rapide diminution des performances des patients implantés lors de la compréhension de parole en condition bruitée.

9.2.3 Influence de la profondeur d'implantation

Plusieurs études psychophysiques réalisées chez le patient implanté cochléaire et par simulation de l'implant chez le sujet sain ont démontré que des distorsions spectrales telles qu'un décalage vers les basses ou les hautes fréquences (Dorman *et al.*, 1997; Fu et Shannon, 1999d), une déformation spectrale non-linéaire (Shannon *et al.*, 1998) ou une compression ou une expansion de la carte fréquentielle (Baskent et Shannon, 2003, 2004) induisent une dégradation de la compréhension de la parole. Par conséquent, cet ensemble d'études suggère qu'une correspondance optimale entre les fréquences de stimulation des électrodes de l'implant et la tonotopie cochléaire devrait être respectée pour optimiser les performances comportementales des patients implantés. Bien que certaines études aient démontré que la compréhension de la parole en cas de distorsion spectrale est susceptible de s'améliorer de façon significative avec l'apprentissage (Rosen *et al.*, 1999; Dorman et Ketten, 2003; Fu *et al.*, 2005), on peut s'attendre à ce qu'une bonne correspondance lieu/fréquence obtenue dès le début permette un bénéfice comportemental optimal pour les patients implantés.

Dans une situation idéale, l'implant cochléaire devrait être inséré dans la rampe tympanique de façon à ce que chaque électrode entre en contact avec la position de la cochlée codant pour la fréquence centrale représentée par cette électrode. En pratique, il est difficile de contrôler précisément cette correspondance entre tonotopie cochléaire et représentation fréquentielle par les électrodes de l'implant. En effet la tonotopie cochléaire varie légèrement d'une personne à l'autre (voir « *Fonctionnement de la cochlée* », § 1.1.2, p. 22) ; par ailleurs, en fonction de la géométrie des conduits cochléaires et de la façon dont l'implant est inséré dans la rampe tympanique, il n'est pas possible de maîtriser complètement le positionnement des électrodes le long de la cochlée.

Afin d'en savoir plus, les recherches dans le domaine s'orientent dans deux directions distinctes : certaines études se donnent pour objectif de vérifier a posteriori la bonne correspondance entre les fréquences représentées par les électrodes de l'implant et la tonotopie cochléaire, tandis que d'autres visent à évaluer les conséquences du décalage spectral induit par un mauvais positionnement, soit chez le patient implanté, soit au travers de simulations acoustiques chez le sujet normo-entendant.

Correspondance entre les électrodes et la tonotopie cochléaire

L'établissement d'une carte de correspondance entre fréquence de stimulation et position dans la cochlée n'est pas trivial. Les cartes lieu/fréquence classiques (Greenwood, 1990) reposent sur l'hypothèse d'une stimulation au niveau des cellules ciliées. Or ces cellules sont la plupart du

temps absentes ou dégradées en cas d'atteinte de l'oreille interne, leur dégradation s'accompagnant d'une dégénération rapide des fibres du nerf radial (Otte *et al.*, 1978; Hinojosa et Marion, 1983; Leake et Hradek, 1988; Nadol *et al.*, 1989; Nadol, 1990; McFadden *et al.*, 2004). Les cibles nerveuses résiduelles sont donc principalement constituées de dendrites et de corps cellulaires dans le ganglion spiral (voir fig. 1.3, p. 21), dont la dégénération se produit beaucoup plus lentement que dans le nerf radial (Fayad *et al.*, 1991; Nadol, 1997; Khan *et al.*, 2005). Toutefois ces cibles nerveuses ne sont pas caractérisées par la même relation lieu/fréquence que les cellules ciliées (Kawano *et al.*, 1996; Sridhar *et al.*, 2006), d'autant plus que pour un même angle d'insertion dans la cochlée, la relation entre les longueurs de l'organe de Corti et du ganglion spiral n'est pas tout à fait linéaire (Stakhovskaya *et al.*, 2007). Par conséquent l'utilisation de la carte tonotopique de Greenwood n'est pas parfaitement adaptée aux cas des patients implantés cochléaires.

La réalisation de cartes de correspondance entre fréquence et lieu de stimulation dans le cas des patients implantés cochléaires nécessite l'examen de patients ayant subi une implantation cochléaire unilatérale et bénéficiant d'une audition résiduelle sur un large domaine fréquentiel dans l'oreille non implantée. Comme les personnes bénéficiant d'une telle audition résiduelle ne sont généralement pas retenues comme candidats à l'implantation cochléaire, de tels cas de figure restent relativement rares. Le première étude dans ce sens a été réalisée par Eddington *et al.* (1978) chez un patient sourd unilatéral, et les relations entre fréquence et position de stimulation se sont trouvées en accord avec les prévisions données par la carte tonotopique de Greenwood ; toutefois les études ultérieures n'ont pas abouti aux mêmes résultats.

Il existe deux façons de repérer la position sur la cochlée, soit par mesure de l'angle d'insertion dans la rampe tympanique, soit par mesure de la distance parcourue à partir de la fenêtre ronde. Quelle que soit la métrique retenue, l'ensemble des études modernes visant à établir la correpondance entre fréquence de stimulation et position dans la cochlée s'accordent pour indiquer, pour une même position cochléaire, un décalage vers les basses fréquences en comparaison avec la carte tonotopique de Greenwood. Ces résultats ont été répétés dans des cas variés avec des patients dont l'oreille non implantée était atteinte d'une surdité sévère (Blamey *et al.*, 1996b; James *et al.*, 2001) ou moyenne (Boex *et al.*, 2006; Dorman *et al.*, 2007), ce dernier cas permettant d'établir des correspondances assez précises entre l'oreille implantée et l'oreille non implantée.

Influence d'un décalage spectral chez le patient implanté

De façon générale, une représentation fréquentielle inadaptée en raison d'un mauvais positionnement de l'implant cochléaire est susceptible de provoquer des difficultés pour s'adapter à la simulation électrique, ainsi qu'une saturation des performances à un bas niveau (Dorman *et al.*, 2007). L'introduction d'un décalage spectral artificiel chez les patients implantés cochléaires est rendue possible par une modification de la programmation du processeur externe, autorisant ainsi des manipulations aussi variées que des décalages, des compressions ou des expansions spectrales. Les études menées dans ce sens s'accordent pour indiquer une plus grande influence d'un décalage spectral sur la perception des indices spectraux que sur la perception des indices temporels, et tendent à montrer de meilleures performances comportementales lorsque le porte-électrode est inséré de façon plus apicale (vers les basses fréquences) (Skinner *et al.*, 1995; Fu et Shannon, 1999b,d).

On dispose de peu de données expérimentales concernant les capacités d'adaptation des patients implantés à un décalage spectral, puisque de telles études demandent aux patients d'utiliser pendant une période prolongée des processeurs fournissant une représentation spectrale déformée de la parole. Toutefois les rares études effectuées dans le domaine démontrent l'existence

d'une adaptation perceptive aux décalages spectraux, permettant de récupérer des niveaux de compréhension similaires à ceux atteints en l'absence de distorsion spectrale (Kileny *et al.*, 1992; Fu *et al.*, 2002; McKay et Henshall, 2002; Dorman et Ketten, 2003)

Influence d'un décalage spectral chez le sujet normo-entendant

Les études utilisant une modification de la programmation du processeur de l'implant chez les patients implantés cochléaires sont complétées par un ensemble d'études parallèles menées chez le sujet normo-entendant grâce à l'utilisation d'une simulation acoustique du traitement effectué par le processeur de l'implant. Cette simulation acoustique est souvent désignée par les termes *vocoder* (Dudley, 1939a,b) ou *SSSN* pour *speech-spectrum shaped noise* (Shannon *et al.*, 1995). La procédure consiste à découper le spectre original de la parole en autant de bandes spectrales qu'il y a d'électrodes à simuler, puis à effectuer dans chaque bande ainsi isolée une uniformisation de l'énergie spectrale sur toutes les fréquences sonores de la bande (cela revient à moduler un bruit blanc de bande par l'amplitude globale de la bande spectrale). Le signal acoustique ainsi obtenu possède de nombreuses propriétés spectrales et temporelles en commun avec le signal électrique délivré à la sortie du processeur de l'implant cochléaire. On pourra consulter *« Simulation acoustique de l'implant cochléaire »*, sect. 10.2.2, p. 218 pour une description détaillée des opérations intervenant dans une simulation d'implant cochléaire.

Comme dans le cas des patients implantés, les études menées dans cette direction de recherche ont démontré la capacité du cerveau à s'adapter rapidement à des déformations spectrales, de façon à récupérer un niveau de compréhension proche de la normale (Dorman *et al.*, 1997; Rosen *et al.*, 1999; Fu *et al.*, 2005; Faulkner, 2006).

9.3 Récupération perceptive : facteurs neurofonctionnels

Une courte durée de déprivation auditive et la présence d'une audition résiduelle comptent parmi les facteurs les mieux corrélés à une bonne compréhension auditive de la parole après une implantation cochléaire (Shallop *et al.*, 1992; Deguine *et al.*, 1993; Blamey *et al.*, 1996a; Kiefer *et al.*, 1998; Rubinstein *et al.*, 1999; van Dijk *et al.*, 1999). L'âge d'implantation et la durée de stimulation par l'implant, le bilan anatomique préopératoire et le taux de survie des cellules des ganglions spiraux comptent également parmi les principaux facteurs associés à l'évolution des performances de compréhension de la parole après implantation (Tyler *et al.*, 1986; Deguine *et al.*, 1993; Waltzman *et al.*, 1995; Blamey *et al.*, 1996a; Fu et Shannon, 1999a).

9.3.1 Facteurs neurophysiologiques

On distingue parmi ces facteurs la proportion de cellules nerveuses intactes (Walsh et Leake-Jones, 1982; Shepherd *et al.*, 1983; Jyung *et al.*, 1989), les propriétés neurophysiologiques du nerf auditif comme sa capacité de retour à la normale après un état réfractaire (Stypulkowski et van den Honert, 1984; Brown *et al.*, 1990, 1996), l'influence de la déprivation sensorielle sur le développement neurophysiologique (Trune, 1982; Lousteau, 1987; Leake *et al.*, 1991), le degré de résolution spatiale et temporelle (Shannon, 1983; Zeng et Shannon, 1994) ainsi que l'intégrité des voies auditives centrales (Stypulkowski *et al.*, 1986; Oviatt et Kileny, 1991; Kraus *et al.*, 1993; Micco *et al.*, 1995).

9.3.2 Effet de l'expérience sensorielle

Âge d'implantation et durée de surdité

Les données de la littérature scientifique montrent que les performances auditives et audio-visuelles après implantation sont d'autant meilleures que les personnes sont implantées jeunes et que la durée de surdité est faible. Ces résultats sont particulièrement marquants dans le cas des enfants implantés cochléaires (Staller *et al.*, 1991; Waltzman *et al.*, 1997; Fryauf-Bertschy *et al.*, 1997; Kirk *et al.*, 2002; Bergeson *et al.*, 2003, 2005; Uziel *et al.*, 2007), mais on peut également observer une bonne corrélation positive entre durée de surdité et récupération post-implantation chez les adultes sourds postlinguaux (UK Cochlear Implant Study Group, 2004). Toutefois la corrélation s'inverse pour les performances visuelles, les durées de surdité plus élevées permettant en effet d'atteindre des scores de compréhension visuelle plus élevés (Bergeson *et al.*, 2003, 2005).

Mode de communication durant la surdité

Plusieurs études montrent l'influence du mode de communication préférentiellement utilisé pendant la période de surdité sur les performances perceptives obtenues à la suite de l'implantation cochléaire. Ainsi, les enfants ayant l'habitude d'utiliser des modes de communication à dominante orale lors de leur surdité (avec l'appui éventuel de la lecture labiale ou du LPC) aboutissent à une meilleure reconnaissance auditive des mots après leur implantation que les enfants utilisant des modes de communication signés (en accompagnement du langage oral) (Hodges *et al.*, 1999; Miyamoto *et al.*, 1999; Cullington *et al.*, 2000; Kirk *et al.*, 2000; Svirsky *et al.*, 2000). Ces résultats ont été étendus par la suite à la compréhension visuelle et audiovisuelle des mots, la communication orale permettant un bénéfice plus important que la communication signée (Lachs *et al.*, 2001; Bergeson *et al.*, 2003).

9.3.3 Période sensible pour l'implantation cochléaire

L'implantation cochléaire chez les adultes sourds prélinguaux permet difficilement de retaurer une compréhension aisée de la parole dans des situations conversationnelles normales (Tyler et Lowder, 1992; Dawson *et al.*, 1992; Busby *et al.*, 1992, 1993; Gantz *et al.*, 1993, 1994). Pour cette raison, on recommande une implantation cochléaire avant l'âge de cinq ans en cas de surdité prélinguale (Fryauf-Bertschy *et al.*, 1997; Waltzman et Cohen, 1998; Schauwers *et al.*, 2004).

Les enfants sourds prélinguaux ayant subi une implantation cochléaire avant l'adolescence présentent un délai dans le développement normal de la morphologie de leurs potentiels corticaux évoqués (Ponton *et al.*, 1996a,b; Eggermont *et al.*, 1997; Gordon *et al.*, 2002). Après stimulation par un implant cochléaire, les latences des potentiels évoqués reprennent leur maturation, tout en conservant un retard de développement incompressible dont la durée correspond à la durée de la surdité. Ceci a conduit les auteurs à faire l'hypothèse que les latences P1 sont réduites pendant le développement par un facteur correspondant à la durée de l'expérience auditive. Toutefois ces études n'incluaient pas les enfants implantés avant l'âge de quatre ans. Des travaux ultérieurs visant à effectuer des mesures similaires sur des enfants implantés avant quatre ans ont permis de montrer que ces derniers rattrapaient leur retard de maturation des latences P1 en quelques mois après l'implantation cochléaire (Sharma *et al.*, 2002a,b,c, 2005). Cet ensemble de résultats met en évidence l'existence d'une période sensible dans le développement des potentiels évoqués auditifs, cette période correspondant à la période sensible d'acquisition de la parole (Fryauf-Bertschy *et al.*, 1997).

Le développement de potentiels N1 maturs est représentatif d'une autre période critique. Les enfants sourds prélinguaux implantés pendant l'adolescence ne développent pas de potentiels N1 maturs, qui suivent les potentiels P1 lors du développement des individus normo-entendants (Ponton et Eggermont, 2001). Ces enfants ne parviennent pas non plus à accéder à une compréhension conversationnelle de la parole sans l'aide de la lecture labiale (Dawson *et al.*, 1992; Tyler et Lowder, 1992; Busby *et al.*, 1992, 1993; Gantz *et al.*, 1993, 1994). Il est connu que les aires auditives de haut niveau ainsi que les aires non-auditives sont recrutées lors du traitement du langage chez les implantés cochléaires, leur activation correspondant à l'accès à une compréhension du langage (Giraud *et al.*, 2000). Par ailleurs les potentiels N1 sont connus pour être engendrés à partir des aires auditives de haut niveau (Liegeois-Chauvel *et al.*, 1994), ce qui laisse penser que ces aires ne sont pas activées de façon appropriée chez les enfants sourds prélinguaux implantés tardivement à l'adolescence.

Ainsi les études des enfants sourds prélinguaux permettent des conclusions approfondies concernant les aspects neurophysiologiques de la plasticité cérébrale dans le système auditif naïf. On peut distinguer au moins deux phases de réorganisation plastique dans le cortex auditif (Kral *et al.*, 2006). La première est une phase extensive et rapide, se mettant en place dans les premières semaines après l'implantation. Les latences P1 maturent rapidement lors de cette première phase plastique chez les enfants sourds prélinguaux ayant subi une implantation cochléaire (Sharma *et al.*, 2005). La seconde phase se met en place plus tard dans le développement, et est plus lente et moins extensive, mais présente également une période sensible.

9.3.4 Utilisation du métabolisme pré-implantation comme facteur prédictif

Si le réseau cortical normalement dédié au traitement auditif du langage se réorganise lors de la surdité en adoptant une spécialisation fonctionnelle irréversible, comme cela semble être le cas pour des surdités prélinguales non traitées, alors la fonction auditive après implantation cochléaire sera contrainte par les propriétés fonctionnelles des régions laissées intactes par la réorganisation plastique, et restant ainsi disponibles pour le traitement auditif (Lee *et al.*, 2007). Si ces régions épargnées se limitent aux aires auditives primaires, alors la perception de la parole pourra être compromise par l'indisponibilité de certaines étapes cruciales dans le traitement de la parole. Inversement, plus les régions cérébrales restant disponibles pour le traitement auditif seront étendues, meilleures seront les chances de récupération du traitement auditif de la parole. Ainsi l'étendue et la localisation de la plasticité cérébrale consécutive à la surdité doivent être précisément évaluées afin de permettre une prédiction qualitative et quantitative de la récupération auditive.

Bien que de nombreuses études montrent que la durée de la surdité permet d'expliquer la majeure partie de la variabilité de l'évolution des patients sourds prélinguaux après une implantation cochléaire (O'Donoghue *et al.*, 2000; Sarant *et al.*, 2001), les observations montrent que d'autres facteurs entrent également en compte. Certaines études apportent ainsi des contre-exemples à un lien direct et trivial entre durée de surdité et récupération post-implantation, en montrant que des enfants sourds prélinguaux implantés plus tardivement présentent parfois une bien meilleure compréhension de la parole que des enfants implantés plus précocemment (Schramm *et al.*, 2002). Jusqu'ici, aucun facteur permettant d'expliquer de façon consistante cette variabilité dans l'évolution de la compréhension de la parole après implantation n'a pu être identifié, bien que certaines variables démographiques ou des facteurs périphériques comme le nombre de cellules préservées dans les ganglions spiraux de la cochlée puissent jouer un rôle mineur dans cette variabilité (Nadol *et al.*, 2001). Toutefois, comme la plasticité corticale varie considérablement d'un individu à un autre, le degré individuel de réorganisation corticale

pourrait constituer un facteur prédictif efficace pour l'évolution du patient après implantation cochléaire (Lee *et al.*, 2001).

Avant l'implantation cochléaire, l'extension de la réorganisation plastique due à la déprivation auditive peut être corrélée au degré d'hypométabolisme dans les lobes temporaux (Lee *et al.*, 2001). Ainsi l'évaluation de l'étendue et de l'intensité de l'hypométabolisme dans ces régions critiques pour la perception de la parole permet d'indiquer si elles restent disponibles pour l'accès au langage auditif, ce qui peut constituer un facteur prédictif intéressant. Si ces aires temporales ne présentent pas d'hypométabolisme, cela indique qu'elles sont impliquées dans des réseaux neurofonctionnels alternatifs et qu'elles ne sont plus disponibles pour une stimulation auditive. Toutefois ce degré d'hypométabolisme est étroitement lié aux aspects temporels de la plasticité corticale (Lee *et al.*, 2001), ce qui ne permet pas de l'utiliser en vue de prédictions indépendantes de l'âge, par exemple pour différencier les évolutions possibles des performances perceptives entre des patients ayant été exposés à des stimulations auditives comparables en durées (Lee *et al.*, 2007).

Plusieurs travaux plus approfondis réalisés chez des enfants sourds prélinguaux ont étudié en détail la corrélation entre le métabolisme cérébral avant implantation et le degré de récupération dans la compréhension de la parole après implantation, en distinguant les facteurs liés à l'âge d'implantation, à la durée de la surdité et aux scores de compréhension de la parole (Oh *et al.*, 2003; Kang *et al.*, 2004; Lee *et al.*, 2005, 2007). Ces travaux ont démontré que le degré de récupération en compréhension de la parole chez les enfants sourds prélinguaux était corrélé au métabolisme pré-implantation. Ils ont en particulier mis en évidence une récupération fonctionnelle d'autant plus élevée que le métabolisme pré-implantation est élevé dans les régions fronto-pariétales (notamment dans le cortex préfrontal gauche) et faible dans les régions temporo-occipitales.

Ces résultats montrent que la récupération fonctionnelle post-implantation tend à être favorisée chez les sujets recrutant en priorité des réseaux cérébraux sous-tendant des fonctions exécutives comme l'attention et la mémoire de travail, situés en particulier dans les cortex préfrontaux et pariétaux, tandis qu'elle peut être plus difficile chez les sujets recrutant en priorité des réseaux cérébraux consacrés à l'analyse perceptive des signaux audiovisuels de parole, situés en particulier dans les régions temporo-occipitales. Cela signifie que lors de la compréhension de la parole chez le sujet implanté cochléaire, l'étape d'analyse audiovisuelle de la parole n'est pas suffisante et qu'elle doit se faire en coopération avec les fonctions cognitives de haut niveau qui sous-tendent la focalisation de l'attention et l'utilisation de la mémoire à court terme, ce qui peut par exemple permettre de tirer parti du contexte sémantique et syntaxique pour mieux comprendre ce qui est dit.

9.4 Réorganisation corticale après implantation cochléaire

9.4.1 Réactivation des voies auditives centrales

Les premières études en neuroimagerie fonctionnelle ont montré que l'utilisation d'un implant cochléaire permettait une réactivation des centres nerveux auditifs primaires et secondaires bilatéraux chez les sourds pré- et postlinguaux, avec une activation préférentielle du côté contralatéral à l'oreille implantée (Herzog *et al.*, 1991; Ito *et al.*, 1993; Truy *et al.*, 1995; Naito *et al.*, 1995; Okazawa *et al.*, 1996). Chez les sourds postlinguaux, on observe une augmentation de l'activité en réponse spécifique aux sons de parole dans les aires auditives associatives et dans les aires frontales inférieures postérieures bilatérales, avec une prédominance de la suractivation dans l'hémisphère gauche (Naito *et al.*, 1995; Okazawa *et al.*, 1996). Plusieurs études ont suggéré

une corrélation entre le degré de métabolisme dans les aires spécifiques au traitement du langage (notamment les aires de Broca et de Wernicke) et les performances de compréhension de la parole chez les sujets sourds implantés cochléaires (Okazawa *et al.*, 1996; Fujiki *et al.*, 1999).

9.4.2 Changements de spécialisation fonctionnelle

Bien que les toutes premières études en neuroimagerie fonctionnelle chez les sujets implantés cochléaires montrent des similarités importantes dans la localisation des aires activées en comparaison avec les sujets normo-entendants, les études ultérieures ont rapidement suggéré l'existence de différences fondamentales entre patients implantés et sujets contrôles lors du traitement cortical de la parole et du langage (Wong *et al.*, 1999; Giraud *et al.*, 2000). Les recherches les plus récentes convergent pour indiquer que pour des niveaux comparables de performances comportementales, les sujets implantés n'utilisent pas les mêmes réseaux corticaux que les sujets normo-entendants.

Une première catégorie de différences est constituée par les changements de spécialisation fonctionnelle : dans ce cas de figure, les différences ne sont pas reflétées par une modification de l'activité cérébrale, mais plutôt par des changements dans les conditions de stimulation dont dépend l'activité cérébrale. En d'autres termes, une région corticale donnée ayant un certain rôle chez le sujet normo-entendant sous-tend une fonction légèrement distincte chez le sujet implanté cochléaire.

Au niveau expérimental, des changements de spécialisation fonctionnelle ont été observés dans le cas de la région temporale supérieure postérieure gauche (aire de Wernicke). Cette région, classiquement impliquée dans le traitement phonologique de la parole chez le sujet normo-entendant, subit après l'implantation cochléaire une forme de déspécialisation et intervient dans le traitement de toutes les catégories sonores susceptibles d'avoir un sens, comme par exemple les sons environnementaux (sonnerie du téléphone, bruits de pas, bruits d'animaux, etc.) (Giraud *et al.*, 2001a).

Ce traitement plus « permissif » de la parole est à mettre en relation avec le ressenti subjectif des patients implantés cochléaires, qui au moment de la mise en marche de l'implant expriment une impression de bruit auditif lorsqu'ils écoutent des sons de parole. Ces données comportementales sont confortées par les études en neuroimagerie fonctionnelle, qui mettent en évidence une activation indifférenciée du cortex auditif aux bruits ou aux sons de parole dans les premiers jours après l'activation de l'implant cochléaire (Giraud *et al.*, 2001c). Ce traitement indifférencié est toutefois transitoire, et le suivi longitudinal des activations corticales montre l'apparition progressive d'un traitement de la parole spécifique et latéralisé à gauche (Giraud, 2005).

Cette spécialisation élargie est cohérente avec les données portant sur la latéralisation fonctionnelle chez le sujet normo-entendant, les sons environnementaux étant dans ce cas traités plus spécifiquement par les régions temporales supérieures droites (Thierry *et al.*, 2003; von Kriegstein et Giraud, 2004). Chez le sujet implanté, au contraire, la distinction entre parole et non-parole est plus difficile à réaliser, et les régions normalement consacrées au traitement phonologique acceptent comme potentiellement phonologiques des sons qui ne le sont pas. Ce type d'activation reflète l'orientation vers une stratégie de traitement entreprenante favorisant les faux-positifs, indiquant qu'il est préférable pour le sujet implanté de soumettre un son de l'environnement à un traitement approfondi de l'hémisphère gauche plutôt que de manquer un son de parole (Giraud, 2005).

9.4.3 Sur- et sous-activations corticales

Une deuxième catégorie de différences séparant patients implantés et sujets normo-entendants lors du traitement de la parole et du langage réside dans les changements des niveaux d'activations corticales. Dans ce cas, la spécialisation fonctionnelle demeure inchangée et seul le niveau d'activation cérébrale augmente ou diminue.

Les principales sous-activations observées chez les sujets implantés se produisent dans les aires temporales ventrales et dans les régions temporo-pariétales (gyrus angulaire), consacrées à l'analyse audiovisuelle et phonologique de la parole (Giraud *et al.*, 2001a). Les sur-activations sont quant à elles observées dans des régions impliquées dans les processus attentionnels et mnésiques, notamment les régions pariétales supérieures et le gyrus parahippocampique (Giraud *et al.*, 2000).

On peut ici faire un parallèle avec les expériences en neuroimagerie où des sujets normo-entendants sont exposés à des sons de parole vocodés (simulation acoustique du traitement effectué par l'implant cochléaire). Ces expériences semblent mettre en évidence des sur- et sous-activations similaires, avec une diminution des traitements dans les régions temporales ventrales et des activations accrues des cortex préfrontaux et pariétaux (Giraud, 2005).

On peut expliquer ces activations différentielles en avançant l'hypothèse que le sujet implanté compense ses difficultés de traitement acoustique et phonologique par des stratégies cognitives de haut niveau qui interagissent en partie avec les étapes de traitement difficiles de bas-niveau via des mécanismes de type « top-down », ce qui permettrait une amélioration du contraste perceptif grâce au contexte sémantique et syntaxique global du flux de parole.

Cette hypothèse est complètement cohérente avec les données de neuroimagerie chez l'enfant implanté cochléaire, où l'on observe chez les sujets obtenant les meilleures performances de reconnaissance de parole un fort métabolisme cérébral pré-implantation dans le cortex pariétal supérieur et l'aire de Broca (Lee *et al.*, 2005) (voir également *« Utilisation du métabolisme pré-implantation comme facteur prédictif »*, § 9.3.4, p. 200). Ces résultats suggèrent que les stratégies cognitives utilisées ont une très grande importance dans les performances globales, en particulier la capacité des sujets à mobiliser leurs ressources attentionnelles et leur mémoire de travail lors de l'écoute de la parole.

9.4.4 Interactions multisensorielles

Compréhension de la parole

Chez le sujet normo-entendant, des activations intermodales des réseaux visuels associatifs par l'audition peuvent être observées lors de l'écoute de la langue maternelle ou d'une langue moins familière mais comprise, tandis qu'elles ne sont plus observées dans le cas d'une langue complètement inconnue. Cette observation suggère que de telles activations intermodales sont fortement dépendantes de processus cognitifs de haut niveau impliqués dans le traitement phonétique ou sémantique de la parole (Giraud et Truy, 2002).

Chez les patients implantés cochléaires, on peut également observer des activations du cortex visuel associatif (BA 18) en l'absence de toute stimulation visuelle, en réponse aux sons de parole (Giraud *et al.*, 2001b; Zatorre, 2001). Ces activations intermodales, inexistantes avant l'implantation cochléaire, se développent au cours des années qui suivent l'implantation ; on observe de plus une spécialisation progressive du cortex visuel dans le traitement des sons de la parole, avec l'apparition de profils d'activation différents pour la parole et les bruits environnementaux (Giraud *et al.*, 2001b).

Les patients implantés présentant les plus fortes activations intermodales du cortex visuel associatif sont également ceux qui possèdent les meilleures performances en lecture labiale (Giraud *et al.*, 2001b). On observe ainsi une progression corrélée entre l'activation du cortex visuel par les sons de parole et les performances en lecture labiale, qui subissent une légère augmentation progressive après implantation (voir également *« Perception visuelle de la parole après implantation »*, § 9.1.2, p. 193). Ce phénomène de plasticité intermodale est différent de la compétition pour l'espace cortical qui se produit pendant la surdité (voir *« Neuroplasticité intermodale après déprivation auditive »*, sect. 8.3, p. 171) ; il s'agit ici d'une coopération entre la vision et l'audition, la restauration de l'audition provoquant le rétablissement des interactions audiovisuelles lors de la compréhension de la parole. Il est donc possible que l'association audiovisuelle des indices phonétiques visuels avec les signaux délivrés par l'implant soit à l'origine de la progression constatée en lecture labiale, celle-ci tirant profit du renforcement des contrastes phonétiques visuels par des contrastes phonétiques auditifs (Giraud, 2005).

Reconnaissance du locuteur

Les observations effectuées au niveau de l'aire fusiforme des visages chez le sujet normo-entendant montrent une activation de cette région lors de l'écoute de voix familières (von Kriegstein *et al.*, 2005, 2006) (voir également *« Activations intermodales »*, § 4.4.1, p. 84), cette activation étant probablement liée à des phénomènes d'imagerie mentale implicite où l'audition de la voix permet l'évocation visuelle du visage. La même observation peut être répétée dans le cas des sujets souffrant d'une prosopagnosie développementale (déficit congénital de la reconnaissance des visages), bien que ces sujets soient incapables de se représenter le visage du locuteur pour l'identifier. Toutefois les sujets prosopagnosiques se montrent moins performants que les sujets normo-entendants dans une tâche purement auditive de reconnaissance de la voix (von Kriegstein *et al.*, 2006), ce qui tend à montrer l'importance de l'accès aux représentations auditives et visuelles pour la reconnaissance du locuteur, même lorsque la stimulation est unimodale.

Dans le cas des patients implantés cochléaires, on observe des activations intermodales plus étendues que chez les sujets normo-entendants, avec notamment un recrutement de l'aire fusiforme des visages y compris lors de la présentation de voix non familières. Giraud (2005) interprète ce résultat en faisant l'hypothèse d'un changement de spécialisation fonctionnelle de l'aire fusiforme chez les patients implantés, cette région étant recrutée afin de permettre un traitement phonologique approfondi de la parole visuelle.

Bibliographie

Baskent, D. et Shannon, R. V. (2003). Speech recognition under conditions of frequency-place compression and expansion. *J. Acoust. Soc. Am.*, 113(4 Pt 1):2064–2076.

Baskent, D. et Shannon, R. V. (2004). Frequency-place compression and expansion in cochlear implant listeners. *J. Acoust. Soc. Am.*, 116(5):3130–3140.

Bergeson, T. R., Pisoni, D. B. et Davis, R. A. (2003). A longitudinal study of audiovisual speech perception by hearing-impaired children with cochlear implants. *Volta Rev.*, 103:347–370.

Bergeson, T. R., Pisoni, D. B. et Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear Hear.*, 26(2):149–164.

Beynon, A. J. et Snik, A. F. (2004). Use of the event-related P300 potential in cochlear implant subjects for the study of strategy-dependent speech processing. *Int. J. Audiol.*, 43:1.

Beynon, A. J., Snik, A. F. et van den Broek, P. (2003). Comparison of different speech coding strategies using a disability-based inventory and speech perception tests in quiet and in noise. *Otol. Neurotol.*, 24(3):392–396.

Blamey, P. J., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., Larky, J., Lindstrom, B., Nedzelski, J., Peterson, A., Shipp, D., Staller, S. et Whitford, L. (1996a). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol. Neurootol.*, 1(5):293–306.

Blamey, P. J., Dooley, G. J., Parisi, E. S. et Clark, G. M. (1996b). Pitch comparisons of acoustically and electrically evoked auditory sensations. *Hear. Res.*, 99(1-2):139–150.

Blamey, P. J., Dowell, R. C., Brown, A. M., Clark, G. M. et Seligman, P. M. (1987). Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors. *J. Acoust. Soc. Am.*, 82(1):48–57.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K. et Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *J. Speech Lang. Hear. Res.*, 44(2):264–285.

Boex, C., Baud, L., Cosendai, G., Sigrist, A., Kos, M. I. et Pelizzone, M. (2006). Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J. Assoc. Res. Otolaryngol.*, 7(2):110–124.

Brown, C. J., Abbas, P. J., Borland, J. et Bertschy, M. R. (1996). Electrically evoked whole nerve action potentials in Ineraid cochlear implant users : responses to different stimulating electrode configurations and comparison to psychophysical responses. *J. Speech Hear. Res.*, 39(3):453–467.

Brown, C. J., Abbas, P. J. et Gantz, B. (1990). Electrically evoked whole-nerve action potentials : data from human cochlear implant users. *J. Acoust. Soc. Am.*, 88(3):1385–1391.

Busby, P. A., Tong, Y. C. et Clark, G. M. (1992). Psychophysical studies using a multiple-electrode cochlear implant in patients who were deafened early in life. *Audiology*, 31(2):95–111.

Busby, P. A., Tong, Y. C. et Clark, G. M. (1993). Electrode position, repetition rate, and speech perception by early- and late-deafened cochlear implant patients. *J. Acoust. Soc. Am.*, 93(2):1058–1067.

Calmels, M. N., Saliba, I., Wanna, G., Cochard, N., Fillaux, J., Deguine, O. et Fraysse, B. (2004). Speech perception and speech intelligibility in children after cochlear implantation. *Int. J. Pediatr. Otorhinolaryngol.*, 68(3):347–351.

Chouard, C. H., Fugain, C., Meyer, B. et Lacombe, H. (1983). Long-term results of the multichannel cochlear implant. *Ann. N.Y. Acad. Sci.*, 405:387–411.

Cullington, H., Hodges, A. V., Butts, S. L., Dolan-Ash, S. et Balkany, T. J. (2000). Comparison of language ability in children with cochlear implants placed in oral and total communication educational settings. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 185:121–123.

Danhauer, J. L., Erratt, J. D. et Edgerton, B. J. (1986). Auditory/visual speech perception by cochlear implant users and hearing aid wearers. *Am. J. Otol.*, 7(5):354–360.

David, E. E., Ostroff, J. M., Shipp, D., Nedzelski, J. M., Chen, J. M., Parnes, L. S., Zimmerman, K., Schramm, D. et Seguin, C. (2003). Speech coding strategies and revised cochlear implant candidacy : an analysis of post-implant performance. *Otol. Neurotol.*, 24(2):228–233.

Dawson, P. W., Blamey, P. J., Rowland, L. C., Dettman, S. J., Clark, G. M., Busby, P. A., Brown, A. M., Dowell, R. C. et Rickards, F. W. (1992). Cochlear implants in children, adolescents, and prelinguistically deafened adults : speech perception. *J. Speech Hear. Res.*, 35(2):401–417.

Deguine, O., Fraysse, B., Uziel, A., Cochard, N., Reuillard-Artieres, F., Piron, J. P. et Mondain, M. (1993). Predictive factors in cochlear implant surgery. *Adv. Otorhinolaryngol.*, 48:142–145.

Dorman, M. F. et Ketten, D. (2003). Adaptation by a cochlear-implant patient to upward shifts in the frequency representation of speech. *Ear Hear.*, 24(5):457–460.

Dorman, M. F., Loizou, P. C. et Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *J. Acoust. Soc. Am.*, 102(5 Pt 1):2993–2996.

Dorman, M. F., Soli, S., Dankowski, K., Smith, L. M., McCandless, G. et Parkin, J. (1990). Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant. *J. Acoust. Soc. Am.*, 88(5):2074–2079.

Dorman, M. F., Spahr, T., Gifford, R., Loiselle, L., McKarns, S., Holden, T., Skinner, M. et Finley, C. (2007). An Electric Frequency-to-place Map for a Cochlear Implant Patient with Hearing in the Nonimplanted Ear. *J. Assoc. Res. Otolaryngol.*, 8(2):234–240.

Dowell, R. C., Martin, L. F., Tong, Y. C., Clark, G. M., Seligman, P. M. et Patrick, J. F. (1982). A 12-consonant study on a multiple-channel cochlear implant patient. *J. Speech Hear. Res.*, 25(4):509–516.

Doyle, K. J., Mills, D., Larky, J., Kessler, D., Luxford, W. M. et Schindler, R. A. (1995). Consonant perception by users of Nucleus and Clarion multichannel cochlear implants. *Am. J. Otol.*, 16(5):676–681.

Dudley, H. (1939a). Remaking Speech. *J. Acoust. Soc. Am.*, 11(2):169–177.

Dudley, H. (1939b). The vocoder. *In Bell Labs Record*, pages 122–126.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G. et Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *Ann. Otol. Rhinol. Laryngol.*, 87(6 Pt 2):1–39.

Eggermont, J. J., Ponton, C. W., Don, M., Waring, M. D. et Kwong, B. (1997). Maturational delays in cortical evoked potentials in cochlear implant users. *Acta Otolaryngol.*, 117(2):161–163.

Faulkner, A. (2006). Adaptation to distorted frequency-to-place maps : implications of simulations in normal listeners for cochlear implants and electroacoustic stimulation. *Audiol. Neurootol.*, 11 Suppl 1:21–26.

Fayad, J., Linthicum, Jr, F. H., Otto, S. R., Galey, F. R. et House, W. F. (1991). Cochlear implants : histopathologic findings related to performance in 16 human temporal bones. *Ann. Otol. Rhinol. Laryngol.*, 100(10):807–811.

Fishman, K. E., Shannon, R. V. et Slattery, W. H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J. Speech Lang. Hear. Res.*, 40(5):1201–1215.

Friesen, L. M., Shannon, R. V., Baskent, D. et Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels : comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.*, 110(2):1150–1163.

Friesen, L. M., Shannon, R. V. et Slattery, 3rd, W. H. (1999). The effect of frequency allocation on phoneme recognition with the Nucleus 22 cochlear implant. *Am. J. Otol.*, 20(6):729–734.

Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D. M., Gantz, B. J. et Woodworth, G. G. (1997). Cochlear implant use by prelingually deafened children : the influences of age at implant and length of device use. *J. Speech Lang. Hear. Res.*, 40(1):183–199.

Fu, Q. J., Nogaki, G. et Galvin, 3rd, J. J. (2005). Auditory training with spectrally shifted speech : implications for cochlear implant patient auditory rehabilitation. *J. Assoc. Res. Otolaryngol.*, 6(2):180–189.

Fu, Q. J. et Shannon, R. V. (1999a). Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users. *J. Acoust. Soc. Am.*, 106(6):65–70.

Fu, Q. J. et Shannon, R. V. (1999b). Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant. *Ear Hear.*, 20(4):332–344.

Fu, Q. J. et Shannon, R. V. (1999c). Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear Hear.*, 20(4):321–331.

Fu, Q. J. et Shannon, R. V. (1999d). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J. Acoust. Soc. Am.*, 105(3):1889–1900.

Fu, Q. J. et Shannon, R. V. (2000). Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners. *J. Acoust. Soc. Am.*, 107(1):589–597.

Fu, Q. J., Shannon, R. V. et Galvin, 3rd, J. J. (2002). Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant. *J. Acoust. Soc. Am.*, 112(4):1664–1674.

Fu, Q. J., Shannon, R. V. et Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition : acoustic and electric hearing. *J. Acoust. Soc. Am.*, 104(6):3586–3596.

Fujiki, N., Naito, Y., Hirano, S., Kojima, H., Shiomi, Y., Nishizawa, S., Konishi, J. et Honjo, I. (1999). Correlation between rCBF and speech perception in cochlear implant users. *Auris Nasus Larynx*, 26(3):229–236.

Fullgrabe, C., Berthommier, F. et Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.*, 211(1-2):74–84.

Gantz, B. J., Tyler, R. S., Woodworth, G. G., Tye-Murray, N. et Fryauf-Bertschy, H. (1994). Results of multichannel cochlear implants in congenital and acquired prelingual deafness in children : five-year follow-up. *Am. J. Otol.*, 15:2.

Gantz, B. J., Woodworth, G. G., Knutson, J. F., Abbas, P. J. et Tyler, R. S. (1993). Multivariate predictors of audiological success with multichannel cochlear implants. *Ann. Otol. Rhinol. Laryngol.*, 102(12):909–916.

Geers, A. (1994). Techniques for assessing auditory speech perception and lipreading enhancement in young deaf children. *Volta Rev.*, 96:85.

Geers, A. et Brenner, C. (1994). Speech perception results : Audition and lipreading enhancement. *Volta Rev.*, 96:97.

Geers, A., Brenner, C. et Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear Hear.*, 24(1 Suppl):24S–35S.

Giraud, A.-L. (2005). Quand la langue maternelle devient étrangère : comprendre la parole avec un implant cochléaire. *In Rencontre IPSEN en ORL*.

Giraud, A.-L., Price, C. J., Graham, J. M. et Frackowiak, R. S. (2001a). Functional plasticity of language-related brain areas after cochlear implantation. *Brain*, 124(Pt 7):1307–1316.

Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E. et Frackowiak, R. S. (2001b). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, 30(3):657–663.

Giraud, A.-L. et Truy, E. (2002). The contribution of visual areas to speech comprehension : a PET study in cochlear implants patients and normal-hearing subjects. *Neuropsychologia*, 40(9):1562–1569.

Giraud, A.-L., Truy, E. et Frackowiak, R. S. (2001c). Imaging plasticity in cochlear implant patients. *Audiol. Neurootol.*, 6(6):381–393.

Giraud, A.-L., Truy, E., Frackowiak, R. S., Gregoire, M. C., Pujol, J. F. et Collet, L. (2000). Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. *Brain*, 123:7.

Goh, W. D., Pisoni, D. B., Kirk, K. I. et Remez, R. E. (2001). Audio-visual perception of sinewave speech in an adult cochlear implant user : a case study. *Ear Hear.*, 22(5):412–419.

Gordon, K. A., Papsin, B. C. et Harrison, R. V. (2002). Auditory brain stem and midbrain development after cochlear implantation in children. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 189:32–37.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species - 29 years later. *J. Acoust. Soc. Am.*, 87:2592–2605.

Hamzavi, J., Baumgartner, W. D., Pok, S. M., Franz, P. et Gstoettner, W. (2003). Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Otolaryngol.*, 123(4):493–498.

Helms, J., Muller, J., Schon, F., Moser, L., Arnold, W., Janssen, T., Ramsden, R., von Ilberg, C., Kiefer, J., Pfennigdorf, T., Gstottner, W., Baumgartner, W., Ehrenberger, K., Skarzynski, H., Ribari, O., Thumfart, W., Stephan, K., Mann, W., Heinemann, M., Zorowka, P., Lippert, K. L., Zenner, H. P., Bohndord, M., Huttenbrink, K. et Hochmair-Desoyer, I. (1997). Evaluation of performance with the COMBI40 cochlear implant in adults : a multicentric clinical study. *ORL J. Otorhinolaryngol. Relat. Spec.*, 59(1):23–35.

Herzog, H., Lamprecht, A., Kuhn, A., Roden, W., Vosteen, K. H. et Feinendegen, L. E. (1991). Cortical activation in profoundly deaf patients during cochlear implant stimulation demonstrated by H_2O^{15} PET. *J. Comput. Assist. Tomogr.*, 15(3):369–375.

Hinojosa, R. et Marion, M. (1983). Histopathology of profound sensorineural deafness. *Ann. N.Y. Acad. Sci.*, 405:459–484.

Hodges, A. V., Dolan Ash, M., Balkany, T. J., Schloffman, J. J. et Butts, S. L. (1999). Speech perception results in children with cochlear implants : contributing factors. *Otolaryngol. Head Neck Surg.*, 121(1):31–34.

Ito, J., Sakakibara, J., Iwasaki, Y. et Yonekura, Y. (1993). Positron emission tomography of auditory sensation in deaf patients and patients with cochlear implants. *Ann. Otol. Rhinol. Laryngol.*, 102(10):797–801.

James, C., Blamey, P. J., Shallop, J. K., Incerti, P. V. et Nicholas, A. M. (2001). Contralateral masking in cochlear implant users with residual hearing in the non-implanted ear. *Audiol. Neurootol.*, 6(2):87–97.

Joris, P. X. et Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J. Acoust. Soc. Am.*, 91(1):215–232.

Jyung, R. W., Miller, J. M. et Cannon, S. C. (1989). Evaluation of eighth nerve integrity by the electrically evoked middle latency response. *Otolaryngol. Head Neck Surg*, 101(6):670–682.

Kaiser, A. R., Kirk, K. I., Lachs, L. et Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J. Speech Lang. Hear. Res.*, 46(2):390–404.

Kang, E., Lee, D. S., Kang, H., Lee, J. S., Oh, S. H., Lee, M. C. et Kim, C. S. (2004). Neural changes associated with speech learning in deaf children following cochlear implantation. *Neuroimage*, 22(3):1173–1181.

Kawano, A., Seldon, H. L. et Clark, G. M. (1996). Computer-aided three-dimensional reconstruction in human cochlear maps : measurement of the lengths of organ of Corti, outer wall, inner wall, and Rosenthal's canal. *Ann. Otol. Rhinol. Laryngol.*, 105(9):701–709.

Khan, A. M., Handzel, O., Damian, D., Eddington, D. K. et Nadol, Jr, J. B. (2005). Effect of cochlear implantation on residual spiral ganglion cell count as determined by comparison with the contralateral nonimplanted inner ear in humans. *Ann. Otol. Rhinol. Laryngol.*, 114(5):381–385.

Kiefer, J., Hohl, S., Sturzebecher, E., Pfennigdorff, T. et Gstoettner, W. (2001). Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the Nucleus CI 24M cochlear implant system. *Audiology*, 40(1):32–42.

Kiefer, J., von Ilberg, C., Reimer, B., Knecht, R., Gall, V., Diller, G., Sturzebecher, E., Pfennigdorff, T. et Spelsberg, A. (1998). Results of cochlear implantation in patients with severe to profound hearing loss–implications for patient selection. *Audiology*, 37(6):382–395.

Kiefer, J., von Ilberg, C., Rupprecht, V., Hubner-Egner, J. et Knecht, R. (2000). Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants : effect of variations in stimulation rate and number of channels. *Ann. Otol. Rhinol. Laryngol.*, 109(11):1009–1020.

Kileny, P. R., Zimmerman-Phillips, S., Zwolan, T. A. et Kemink, J. L. (1992). Effects of channel number and place of stimulation on performance with the Cochlear Corporation multichannel implant. *Am. J. Otol.*, 13(2):117–123.

Kirk, K. I., Miyamoto, R. T., Ying, E. A., Perdew, A. E. et Zuganelis, H. (2002). Cochlear implantation in young children : Effects of age at implantation and communication mode. *Volta Rev.*, 102:127.

Kirk, K. I., Pisoni, D. B. et Miyamoto, R. T. (2000). Lexical discrimination by children with cochlear implants : Effects of age at implantation and communication mode. *In* Waltzman, S. et Cohen, N., éditeurs : *Proceedings of the Vth International Cochlear Implant Conference*. New York : Thieme Medical Publishers.

Kompis, M., Vischer, M. W. et Hausler, R. (1999). Performance of compressed analogue (CA) and continuous interleaved sampling (CIS) coding strategies for cochlear implants in quiet and noise. *Acta Otolaryngol.*, 119(6):659–664.

Kral, A., Tillein, J., Heid, S., Klinke, R. et Hartmann, R. (2006). Cochlear implants : cortical plasticity in congenital deprivation. *Prog. Brain Res.*, 157:283–313.

Kraus, N., Micco, A. G., Koch, D. B., McGee, T., Carrell, T., Sharma, A., Wiet, R. J. et Weingarten, C. Z. (1993). The mismatch negativity cortical evoked potential elicited by speech in cochlear-implant users. *Hear. Res.*, 65(1-2):118–124.

Lachs, L., Pisoni, D. B. et Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants : a first report. *Ear Hear.*, 22(3):236–251.

Leake, P. A. et Hradek, G. T. (1988). Cochlear pathology of long term neomycin induced deafness in cats. *Hear. Res.*, 33(1):11–33.

Leake, P. A., Hradek, G. T., Rebscher, S. J. et Snyder, R. L. (1991). Chronic intracochlear electrical stimulation induces selective survival of spiral ganglion neurons in neonatally deafened cats. *Hear. Res.*, 54(2):251–271.

Lee, D. S., Lee, J. S., Oh, S. H., Kim, S. K., Kim, J. W., Chung, J. K., Lee, M. C. et Kim, C. S. (2001). Cross-modal plasticity and cochlear implants. *Nature*, 409(6817):149–150.

Lee, H. J., Giraud, A.-L., Kang, E., Oh, S. H., Kang, H., Kim, C. S. et Lee, D. S. (2007). Cortical activity at rest predicts cochlear implantation outcome. *Cereb. Cortex*, 17(4):909–917.

Lee, H. J., Kang, E., Oh, S. H., Kang, H., Lee, D. S., Lee, M. C. et Kim, C. S. (2005). Preoperative differences of cerebral metabolism relate to the outcome of cochlear implants in congenitally deaf children. *Hear. Res.*, 203:2–9.

Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P. et Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man : evaluation and topography of the middle latency components. *Electroencephalogr. Clin. Neurophysiol.*, 92(3):204–214.

Loizou, P. C., Dorman, M. et Tu, Z. (1999). On the number of channels needed to understand speech. *J. Acoust. Soc. Am.*, 106(4 Pt 1):2097–2103.

Loizou, P. C., Poroy, O. et Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *J. Acoust. Soc. Am.*, 108(2):790–802.

Loizou, P. C., Stickney, G., Mishra, L. et Assmann, P. (2003). Comparison of speech processing strategies used in the Clarion implant processor. *Ear Hear.*, 24(1):12–19.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S. et Moore, B. C. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U.S.A.*, 103(49):18866–18869.

Lousteau, R. J. (1987). Increased spiral ganglion cell survival in electrically stimulated, deafened guinea pig cochleae. *Laryngoscope*, 97(7 Pt 1):836–842.

Manrique, M., Huarte, A., Morera, C., Caballe, L., Ramos, A., Castillo, C., Garcia-Ibanez, L., Estrada, E. et Juan, E. (2005). Speech perception with the ACE and the SPEAK speech coding strategies for children implanted with the Nucleus cochlear implant. *Int. J. Pediatr. Otorhinolaryngol.*, 69(12):1667–1674.

McFadden, S. L., Ding, D., Jiang, H. et Salvi, R. J. (2004). Time course of efferent fiber and spiral ganglion cell degeneration following complete hair cell loss in the chinchilla. *Brain Res.*, 997(1):40–51.

McKay, C. M. et Henshall, K. R. (2002). Frequency-to-electrode allocation and speech perception with cochlear implants. *J. Acoust. Soc. Am.*, 111(2):1036–1044.

McKay, C. M. et McDermott, H. J. (1993). Perceptual performance of subjects with cochlear implants using the spectral maxima sound processor (smsp) and the mini speech processor (msp). *Ear Hear.*, 14(5):350–367.

Micco, A. G., Kraus, N., Koch, D. B., McGee, T. J., Carrell, T. D., Sharma, A., Nicol, T. et Wiet, R. J. (1995). Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am. J. Otol.*, 16(4):514–520.

Miyamoto, R. T., Kirk, K. I., Svirsky, M. A. et Sehgal, S. T. (1999). Communication skills in pediatric cochlear implant recipients. *Acta Otolaryngol.*, 119(2):219–224.

Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer, Jr., E. T., Fisher, L. et Bernstein, L. E. (2005). Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol. Neurotol.*, 26(4): 649–654.

Munson, B. et Nelson, P. B. (2005). Phonetic identification in quiet and in noise by listeners with cochlear implants. *J. Acoust. Soc. Am.*, 118(4):2607–2617.

Nadol, Jr, J. B. (1990). Degeneration of cochlear neurons as seen in the spiral ganglion of man. *Hear. Res.*, 49(1-3):141–154.

Nadol, Jr, J. B. (1997). Patterns of neural degeneration in the human cochlea and auditory nerve : implications for cochlear implantation. *Otolaryngol. Head Neck Surg.*, 117(3 Pt 1):220–228.

Nadol, Jr., J. B., Shiao, J. Y., Burgess, B. J., Ketten, D. R., Eddington, D. K., Gantz, B. J., Kos, I., Montandon, P., Coker, N. J., Roland, Jr., J. T. et Shallop, J. K. (2001). Histopathology of cochlear implants in humans. *Ann. Otol. Rhinol. Laryngol.*, 110:883–891.

Nadol, Jr, J. B., Young, Y. S. et Glynn, R. J. (1989). Survival of spiral ganglion cells in profound sensorineural hearing loss : implications for cochlear implantation. *Ann. Otol. Rhinol. Laryngol.*, 98(6):411–416.

Naito, Y., Okazawa, H., Honjo, I., Hirano, S., Takahashi, H., Shiomi, Y., Hoji, W., Kawano, M., Ishizu, K. et Yonekura, Y. (1995). Cortical activation with sound stimulation in cochlear implant users demonstrated by positron emission tomography. *Cogn. Brain Res.*, 2(3):207–214.

Nelson, P. B., Jin, S. H., Carney, A. E. et Nelson, D. A. (2003). Understanding speech in modulated interference : cochlear implant users and normal-hearing listeners. *J. Acoust. Soc. Am.*, 113(2):961–968.

O'Donoghue, G. M., Nikolopoulos, T. P. et Archbold, S. M. (2000). Determinants of speech perception in children after cochlear implantation. *Lancet*, 356:466–468.

Oh, S. H., Kim, C. S., Kang, E. J., Lee, D. S., Lee, H. J., Chang, S. O., Ahn, S. H., Hwang, C. H., Park, H. J. et Koo, J. W. (2003). Speech perception after cochlear implantation over a 4-year time period. *Acta Otolaryngol.*, 123(2):148–153.

Okazawa, H., Naito, Y., Yonekura, Y., Sadato, N., Hirano, S., Nishizawa, S., Magata, Y., Ishizu, K., Tamaki, N., Honjo, I. et Konishi, J. (1996). Cochlear implant efficiency in pre- and postlingually deaf subjects. A study with H_2O^{15} and PET. *Brain*, 119:4.

Otte, J., Schunknecht, H. F. et Kerr, A. G. (1978). Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *Laryngoscope*, 88(8 Pt 1):1231–1246.

Oviatt, D. L. et Kileny, P. R. (1991). Auditory event-related potentials elicited from cochlear implant recipients and hearing subjects. *Am. J. Audiol.*, 1:48–55.

Parkinson, A. J., Parkinson, W. S., Tyler, R. S., Lowder, M. W. et Gantz, B. J. (1998). Speech perception performance in experienced cochlear-implant patients receiving the SPEAK processing strategy in the Nucleus Spectra-22 cochlear implant. *J. Speech Lang. Hear. Res.*, 41(5):1073–1087.

Pelizzone, M., Cosendai, G. et Tinembart, J. (1999). Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for Ineraid implanted subjects. *Ear Hear.*, 20(3):228–237.

Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., Kwong, B. et Masuda, A. (1996a). Auditory system plasticity in children after long periods of complete deafness. *Neuroreport*, 8(1):61–65.

Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D. et Masuda, A. (1996b). Maturation of human cortical auditory function : differences between normal-hearing children and children with cochlear implants. *Ear Hear.*, 17(5):430–437.

Ponton, C. W. et Eggermont, J. J. (2001). Of kittens and kids : altered cortical maturation following profound deafness and cochlear implant use. *Audiol. Neurootol.*, 6(6):363–380.

Qin, M. K. et Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.*, 114(1):446–454.

Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A. et Cuneo, P. A. (1992). Relations among different measures of speech reception in subjects using a cochlear implant. *J. Acoust. Soc. Am.*, 92(4 Pt 1):1869–1881.

Rose, J. E., Brugge, J. F., Anderson, D. J. et Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.*, 30:769–793.

Rosen, S. (1992). Temporal information in speech : acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 336(1278):367–373.

Rosen, S., Faulkner, A. et Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech : implications for cochlear implants. *J. Acoust. Soc. Am.*, 106(6):3629–3636.

Rosen, S., Walliker, J., Brimacombe, J. A. et Edgerton, B. J. (1989). Prosodic and segmental aspects of speech perception with the house/3m single-channel implant. *J. Speech Hear. Res.*, 32(1):93–111.

Rubinstein, J. T., Parkinson, W. S., Tyler, R. S. et Gantz, B. J. (1999). Residual speech recognition and cochlear implant performance : effects of implantation criteria. *Am. J. Otol.*, 20(4):445–452.

Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M. et Gibson, W. P. (2001). Variation in speech perception scores among children with cochlear implants. *Ear Hear.*, 22:18–28.

Schauwers, K., Gillis, S., Daemers, K., De Beukelaer, C. et Govaerts, P. J. (2004). Cochlear implantation between 5 and 20 months of age : the onset of babbling and the audiologic outcome. *Otol. Neurotol.*, 25(3):263–270.

Schramm, D., Fitzpatrick, E. et Seguin, C. (2002). Cochlear implantation for adolescents and adults with prelinguistic deafness. *Otol. Neurotol.*, 23:698–703.

Shallop, J. K., Arndt, P. L. et Turnacliff, K. A. (1992). Expanded indications for cochlear implantation : Perceptual results in seven adults with residual hearing. *J. Speech Lang. Path.*, 16:141–148.

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. basic psychophysics. *Hear. Res.*, 11(2):157–189.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonsky, J. et Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270:303–304.

Shannon, R. V., Zeng, F. G. et Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *J. Acoust. Soc. Am.*, 104(4):2467–2476.

Sharma, A., Dorman, M., Spahr, A. et Todd, N. W. (2002a). Early cochlear implantation in children allows normal development of central auditory pathways. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 189:38–41.

Sharma, A., Dorman, M. F. et Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hear. Res.*, 203(1-2):134–143.

Sharma, A., Dorman, M. F. et Spahr, A. J. (2002b). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *Neuroreport*, 13(10):1365–1368.

Sharma, A., Dorman, M. F. et Spahr, A. J. (2002c). A sensitive period for the development of the central auditory system in children with cochlear implants : implications for age of implantation. *Ear Hear.*, 23(6):532–539.

Shepherd, R. K., Clark, G. M. et Black, R. C. (1983). Chronic electrical stimulation of the auditory nerve in cats. physiological and histopathological results. *Acta Otolaryngol. Suppl.*, 399:19–31.

Skinner, M. W., Fourakis, M. S., Holden, T. A., Holden, L. K. et Demorest, M. E. (1999). Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies. II. consonants. *Ear Hear.*, 20(6):443–460.

Skinner, M. W., Holden, L. K. et Holden, T. A. (1995). Effect of frequency boundary assignment on speech recognition with the speak speech-coding strategy. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 166:307–311.

Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C. et Holden, T. A. (2002). Speech recognition with the Nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear Hear.*, 23(3):207–223.

Smith, Z. M., Delgutte, B. et Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876):87–90.

Souza, P. E. et Boike, K. T. (2006). Combining temporal-envelope cues across channels : effects of age and hearing loss. *J. Speech Lang. Hear. Res.*, 49(1):138–149.

Sridhar, D., Stakhovskaya, O. et Leake, P. A. (2006). A frequency-position function for the human cochlear spiral ganglion. *Audiol. Neurootol.*, 11 Suppl 1:16–20.

Stakhovskaya, O., Sridhar, D., Bonham, B. H. et Leake, P. A. (2007). Frequency map for the human cochlear spiral ganglion : implications for cochlear implants. *J. Assoc. Res. Otolaryngol.*, 8(2):220–233.

Staller, S. J., Dowell, R. C., Beiter, A. L. et Brimacombe, J. A. (1991). Perceptual abilities of children with the Nucleus 22-channel cochlear implant. *Ear Hear.*, 12(4 Suppl):34S–47S.

Stobich, B., Zierhofer, C. M. et Hochmair, E. S. (1999). Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear Hear.*, 20(2):104–116.

Stone, M. A. et Moore, B. C. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *J. Acoust. Soc. Am.*, 114(2):1023–1034.

Stypulkowski, P. H. et van den Honert, C. (1984). Physiological properties of the electrically stimulated auditory nerve. I. compound action potential recordings. *Hear. Res.*, 14(3):205–223.

Stypulkowski, P. H., van den Honert, C. et Kvistad, S. D. (1986). Electrophysiologic evaluation of the cochlear implant patient. *Otolaryngol. Clin. North Am.*, 19(2):249–257.

Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B. et Miyamoto, R. T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychol. Sci.*, 11(2):153–158.

Thierry, G., Giraud, A.-L. et Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron*, 38(3):499–506.

Trune, D. R. (1982). Influence of neonatal cochlear removal on the development of mouse cochlear nucleus : I. number, size, and density of its neurons. *J. Comp. Neurol.*, 209(4):409–424.

Truy, E., Deiber, M. P., Cinotti, L., Mauguiere, F., Froment, J. C. et Morgon, A. (1995). Auditory cortex activity changes in long-term sensorineural deprivation during crude cochlear electrical stimulation : evaluation by positron emission tomography. *Hear. Res.*, 86(1-2):34–42.

Turner, C. W., Souza, P. E. et Forget, L. N. (1995). Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *J. Acoust. Soc. Am.*, 97(4):2568–2576.

Tyler, R. S., Gantz, B. J., Parkinson, A. J., Woodworth, G. G., Lowder, M. W. et Parkinson, W. S. (1997). Initial comparison between the Clarion, Nucleus and Ineraid cochlear implants. *Adv. Otorhinolaryngol.*, 52:274–278.

Tyler, R. S. et Lowder, M. W. (1992). Audiological management and performance of adult cochlear-implant patients. *Ear Nose Throat J.*, 71(3):117–122, 125–128.

Tyler, R. S., Preece, J. P., Lansing, C. R., Otto, S. R. et Gantz, B. J. (1986). Previous experience as a counfouding factor in comparing cochlear-implant processing schemes. *J. Speech Hear. Res.*, 29(2):282–287.

UK Cochlear Implant Study Group (2004). Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults I : theory and measures of effectiveness. *Ear Hear.*, 25(4):310–335.

Uziel, A. S., Sillon, M., Vieu, A., Artieres, F., Piron, J. P., Daures, J. P. et Mondain, M. (2007). Ten-year follow-up of a consecutive series of children with multichannel cochlear implants. *Otol. Neurotol.*, 28(5):615–628.

Valimaa, T. T. (2002). *Speech perception and auditory performance in hearing-impaired adults with a multichannel cochlear implant*. Thèse de doctorat, Université d'Oulu.

Valimaa, T. T., Maatta, T. K., Lopponen, H. J. et Sorri, M. J. (2002). Phoneme recognition and confusions with multichannel cochlear implants : consonants. *J. Speech Lang. Hear. Res.*, 45(5):1055–1069.

van Dijk, J. E., van Olphen, A. F., Langereis, M. C., Mens, L. H., Brokx, J. P. et Smoorenburg, G. F. (1999). Predictors of cochlear implant performance. *Audiology*, 38(2):109–116.

van Wieringen, A. et Wouters, J. (1999). Natural vowel and consonant recognition by Laura cochlear implantees. *Ear Hear.*, 20(2):89–103.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. et McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients : a comparison of sound-processing strategies. *J. Acoust. Soc. Am.*, 117(5):3126–3138.

von Kriegstein, K. et Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2):948–955.

von Kriegstein, K., Kleinschmidt, A. et Giraud, A.-L. (2006). Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cereb. Cortex*, 16(9):1314–1322.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P. et Giraud, A.-L. (2005). Interaction of face and voice areas during speaker recognition. *J. Cogn. Neurosci.*, 17(3):367–376.

Walsh, S. M. et Leake-Jones, P. A. (1982). Chronic electrical stimulation of auditory nerve in cat : physiological and histological results. *Hear. Res.*, 7(3):281–304.

Waltzman, S. B. et Cohen, N. L. (1998). Cochlear implantation in children younger than 2 years old. *Am. J. Otol.*, 19(2):158–162.

Waltzman, S. B., Cohen, N. L., Gomolin, R. H., Green, J. E., Shapiro, W. H., Hoffman, R. A. et Roland, Jr, J. T. (1997). Open-set speech perception in congenitally deaf children using cochlear implants. *Am. J. Otol.*, 18(3):342–349.

Waltzman, S. B., Fisher, S. G., Niparko, J. K. et Cohen, N. L. (1995). Predictors of postoperative performance with cochlear implants. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 165:15–18.

Whitford, L. A., Seligman, P. M., Everingham, C. E., Antognelli, T., Skok, M. C., Hollow, R. D., Plant, K. L., Gerin, E. S., Staller, S. J., McDermott, H. J., Gibson, W. R. et Clark, G. M. (1995). Evaluation of the Nucleus Spectra 22 processor and new speech processing strategy (SPEAK) in postlinguistically deafened adults. *Acta Otolaryngol.*, 115(5):629–637.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. et Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, 352(6332):236–238.

Wong, D., Miyamoto, R. T., Pisoni, D. B., Sehgal, M. et Hutchins, G. D. (1999). PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. *Hear. Res.*, 132(1-2):34–42.

Zatorre, R. J. (2001). Do you see what I'm saying ? interactions between auditory and visual cortices in cochlear implant users. *Neuron*, 31(1):13–14.

Zeng, F. G. (2004). Trends in cochlear implants. *Trends Amplif.*, 8(1):1–34.

Zeng, F. G. et Galvin, 3rd, J. J. (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners. *Ear Hear.*, 20(1):60–74.

Zeng, F. G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y. Y. et Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues. *J. Acoust. Soc. Am.*, 116(3):1351–1354.

Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C. et Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. U.S.A.*, 102(7):2293–2298.

Zeng, F. G. et Shannon, R. V. (1994). Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. *Science*, 264(5158):564–566.

Ziese, M., Stutzel, A., von Specht, H., Begall, K., Freigang, B., Sroka, S. et Nopp, P. (2000). Speech understanding with the CIS and the n-of-m strategy in the MED-EL COMBI 40+ system. *ORL J. Otorhinolaryngol. Relat. Spec.*, 62(6):321–329.

Zwolan, T. A., Kileny, P. R., Smith, S., Waltzman, S., Chute, P., Domico, E., Firszt, J., Hodges, A., Mills, D., Whearty, M., Osberger, M. J. et Fisher, L. (2005). Comparison of continuous interleaved sampling and simultaneous analog stimulation speech processing strategies in newly implanted adults with a Clarion 1.2 cochlear implant. *Otol. Neurotol.*, 26(3):455–465.

Troisième partie

Etude des conséquences de l'implantation cochléaire

Chapitre 10

Problématique et méthodes

Sommaire

10.1 Problématique 217
10.2 Préparation des stimuli comportementaux 218
10.2.1 Génération des stimuli 218
10.2.2 Simulation acoustique de l'implant cochléaire 218
10.3 Traitement phonétique des réponses comportementales 224
10.3.1 Analyse phonétique des réponses 224
10.3.2 Matrices de confusions phonétiques 225
10.3.3 Analyse de la transmission d'information phonétique 225
10.4 Neuroimagerie fonctionnelle 227
10.4.1 Déroulement d'un examen 227
10.4.2 Dispositif expérimental 227
10.4.3 Prétraitement des images d'activation cérébrale 228
Bibliographie 229

Introduction

Afin d'introduire les différentes études comportementales et neurofonctionnelles que nous avons effectuées, nous récapitulons dans ce chapitre les questions qui restent en suspens concernant la perception audiovisuelle de la parole chez le patient implanté cochléaire, ainsi que les directions de recherche qu'elles suscitent. Nous décrivons également les méthodes expérimentales qui sont communes à nos différentes études.

10.1 Problématique

Après une déprivation sensorielle, quelle que soit la modalité perceptive affectée, on assiste à la réorganisation progressive des mécanismes perceptifs, à la fois au niveau fonctionnel et cortical (voir *« La surdité »*, ch. 6, p. 127 et *« Réorganisation neurofonctionnelle après une altération sensorielle »*, ch. 8, p. 155). Ainsi, dans le cas d'une surdité, la personne privée de son audition va développer des compétences perceptives spécifiques, principalement au niveau de la vision, et

tout particulièrement dans les contextes de communication avec les semblables : l'utilisation de la lecture labiale joue alors un rôle essentiel dans la compréhension de la parole.

Une fois que la personne sourde a subi une implantation cochléaire, ce qui permet la restauration partielle des voies auditives, une nouvelle forme de réorganisation fonctionnelle et corticale va émerger (voir *« Réorganisation neurofonctionnelle après implantation cochléaire »*, ch. 9, p. 191), et on assistera notamment à une progression de la compréhension auditive et audiovisuelle de la parole. On dispose toutefois de très peu de données scientifiques permettant d'évaluer les contributions respectives de la vision et de l'audition dans cette récupération progressive de la communication parlée après une implantation cochléaire, et les mécanismes corticaux associés ne sont pas connus.

Afin d'en savoir plus, nous proposons d'analyser les mécanismes de la perception audiovisuelle de la parole chez des patients sourds postlinguaux implantés cochléaires et chez des sujets contrôles normo-entendants, aussi bien par des méthodes comportementales qu'à travers des protocoles en neuroimagerie.

10.2 Préparation des stimuli comportementaux

10.2.1 Génération des stimuli

Les études psychophysiques et neurofonctionnelles décrites dans la suite utilisent pour la plupart des enregistrements vidéo de prononciations de mots dissyllabiques. Ces mots ont été sélectionnés à partir des listes normalisées de Fournier, utilisées de façon régulière pour les tests orthophoniques en français.

Une orthophoniste du service ORL du CHU Toulouse-Purpan a effectué une prononciation de 110 mots dissyllabiques extraits des listes de Fournier en recevant l'instruction de parler avec régularité, avec une intonation et un niveau normaux. Ces prononciations ont été enregistrées à l'aide d'une caméra numérique professionnelle dans une chambre d'enregistrement aux parois traitées acoustiquement, avec un éclairage uniforme de bonne qualité. Un microphone était ajusté face à la locutrice et le cadrage de la caméra a été réglé de façon à ce que l'image montre de face le visage entier de la locutrice sur un arrière-plan de couleur uniforme.

Les enregistrements ont été numérisés en ligne à une résolution vidéo de 720×576 pixels avec 25 images par seconde. La numérisation audio a été effectuée par une quantification sur 16 bits à 32 kHz. Chaque prononciation a ensuite été isolée numériquement de façon à inclure un court temps de repos avant et après la durée de prononciation.

10.2.2 Simulation acoustique de l'implant cochléaire

Nous décrivons ici en détail les méthodes numériques que nous avons utilisées pour réaliser les simulations acoustiques d'implant cochléaire employées dans nos études chez le sujet normo-entendant (*« Sujet sain : parole traitée par l'implant »*, sect. 11.3, p. 242 et *« Entraînement perceptif avec une simulation d'implant »*, sect. 11.4, p. 249). Ces simulations sont souvent désignées dans la littérature par les appellations *vocoder* (Dudley, 1939a,b) ou *SSSN*, pour *speech-spectrum shaped noise* (Shannon *et al.*, 1995).

Dans un implant cochléaire, chaque électrode effectue un traitement des informations acoustiques associées à une bande spectrale spécifique à cette électrode. Les fréquences perceptibles correspondant à cette bande spectrale vont dépendre de plusieurs facteurs techniques et physiologiques, dont en particulier les réglages spécifiques au processeur de l'implant, le positionnement de l'électrode le long de la rampe tympanique ainsi que l'étendue des fibres nerveuses

du ganglion spiral qui seront stimulées électriquement par cette électrode. Dans l'optique d'une simulation informatique du traitement effectué par l'implant, il sera donc nécessaire de traiter de façon séparée les différents canaux fréquentiels correspondant aux différentes électrodes de l'implant cochléaire. Par conséquent une simulation du traitement acoustique effectué par l'implant va se structurer en fonction des étapes suivantes (voir fig. 10.1, p. 220) :

1. définition des limites spectrales des différents canaux fréquentiels ;
2. décomposition du signal en sous-signaux associés à chaque canal fréquentiel ;
3. éventuels traitements spécifiques à chaque bande spectrale ;
4. recomposition du signal à partir des différents sous-signaux.

Sélection des fréquences de coupure

Les fréquences de coupure qui délimitent les différents canaux fréquentiels correspondent théoriquement aux fréquences codées par les points de la cochlée situés à mi-chemin entre deux électrodes successives. Dans le cadre des simulations acoustiques utilisées dans la littérature, on se base systématiquement sur l'hypothèse d'une répartition équidistante des électrodes le long de la rampe tympanique. Cette hypothèse est cohérente avec la structure des implants Nucleus, MedEl et Ineraid, dont l'espacement inter-électrodes est constant, mais n'est pas adaptée au cas des implants Clarion, pour lesquels l'espacement inter-électrodes se réduit de la base vers l'apex (voir *« Electrodes »*, § 7.5.2, p. 151). Comme les études que nous avons effectuées se limitent exclusivement aux implants Nucleus, l'hypothèse d'équidistance des électrodes le long de la cochlée est ici parfaitement justifiée.

Afin de disposer d'une correspondance tonotopique entre la profondeur d'insertion dans la rampe tympanique et la fréquence sonore codée, les simulations acoustiques d'implant utilisées dans la littérature se basent toutes sur la carte tonotopique de Greenwood (Greenwood, 1990). Celle-ci présente des formules numériques reliant explicitement la position sur la cochlée et la fréquence codée. Ainsi, pour une profondeur d'insertion x exprimée en mm (mesurée à partir de l'apex et non pas de la base) et une fréquence perçue f exprimée en Hz, les relations entre f et x sont les suivantes :

$$f = 165.4 \times (10^{0.06x} - 1)$$

$$x = \frac{1}{0.06} \log \frac{f}{165.4} + 1$$

Cette carte de correspondance entre position et fréquence est toutefois à considérer avec précaution. Elle est en effet basée sur l'hypothèse d'une stimulation électrique au niveau des cellules ciliées, tandis que ces cellules sont la plupart du temps dégradées chez les patients implantés, de sorte que la stimulation électrique se produit en réalité au niveau des cellules du ganglion spiral (voir *« Influence de la profondeur d'implantation »*, § 9.2.3, p. 196). Or la longueur d'enroulement de la cochlée, lorsqu'elle est exprimée en fonction de l'angle d'insertion, varie de façon différente dans la rampe tympanique et dans le ganglion spiral (Stakhovskaya *et al.*, 2007), ce qui suggère un décalage entre la fréquence théorique indiquée par la carte de Greenwood et la fréquence effective observée chez les patients implantés cochléaires. Un tel décalage a été mesuré avec une bonne précision chez des patients implantés unilatéraux bénéficiant de restes auditifs dans l'oreille non implantée (Blamey *et al.*, 1996; James *et al.*, 2001; Boex *et al.*, 2006; Dorman *et al.*, 2007), indiquant un biais vers les basses fréquences pour la fréquence effectivement perçue (voir également *« Influence de la profondeur d'implantation »*, § 9.2.3, p. 196). Toutefois, comme ce biais est

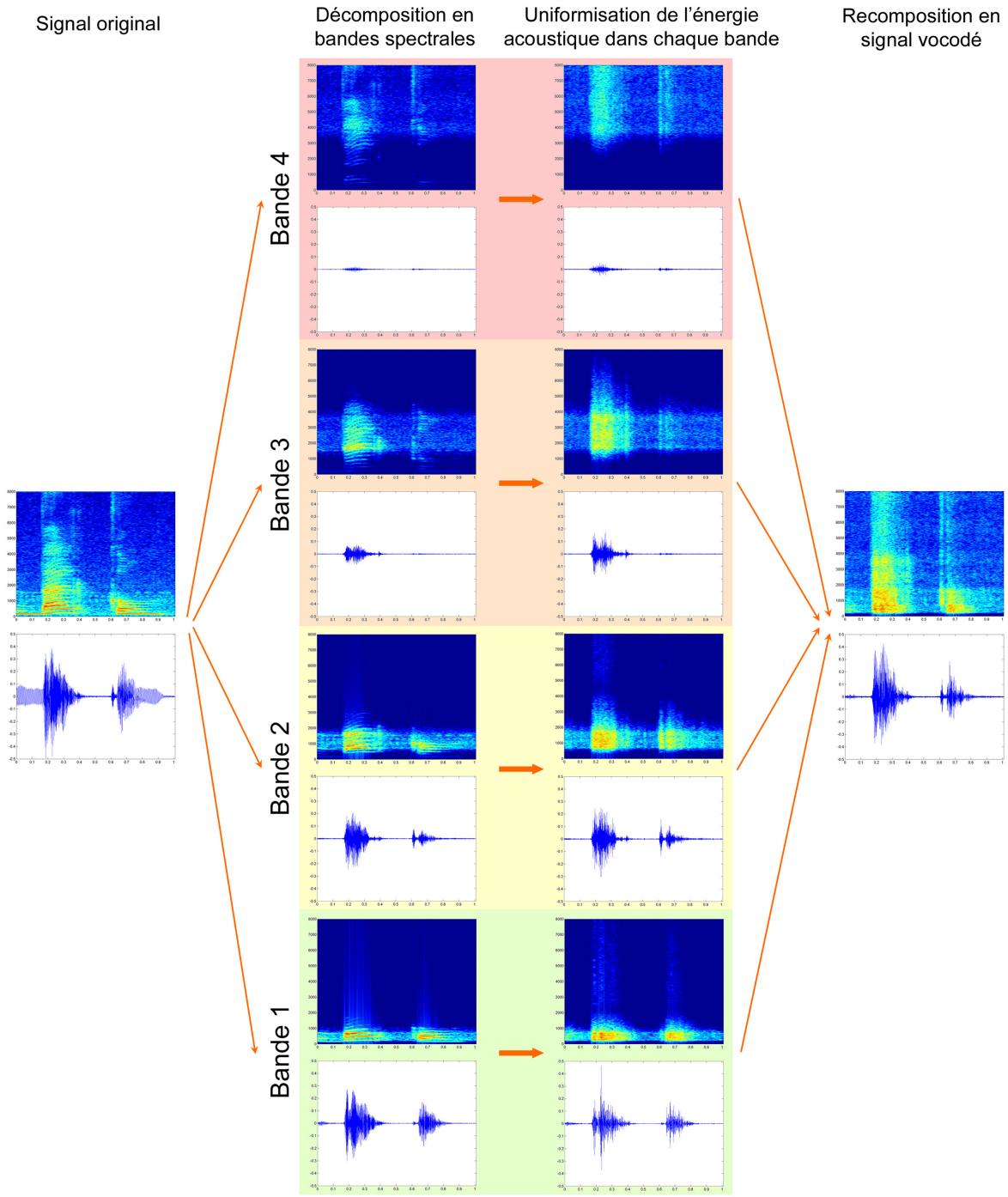


FIG. 10.1 – Etapes de transformation du signal acoustique lors d'une simulation numérique de l'implant cochléaire. Le mot analysé est ici *balcon*, avec l'exemple d'une décomposition suivant quatre bandes fréquentielles. Chaque signal acoustique est représenté sous forme de courbe temps/amplitude et de spectrogramme temps/fréquence. Pour les spectrogrammes, l'énergie acoustique est représentée sur une échelle de couleurs logarithmique, ce qui fausse l'évaluation visuelle de la quantité totale d'énergie contenue dans une bande spectrale.

particulièrement variable en fonction de la physiologie individuelle des structures cochléaires et nerveuses, si bien qu'aucune carte tonotopique adaptée au cas des implantés cochléaire n'existe à l'heure actuelle, nous resterons dans le cadre de la carte tonotopique de Greenwood pour définir nos simulations acoustiques d'implant.

Comme nous nous sommes volontairement placés dans le cadre d'une équidistance des électrodes le long de la cochlée, il ne nous reste plus qu'à choisir la position de la première et de la dernière électrode afin d'en déduire les positions de chacune des électrodes. Ensuite, une fois connues les positions de toutes les électrodes, nous pourrons utiliser la carte de correspondance tonotopique de Greenwood pour en déduire les fréquences de coupure correspondantes. Afin de disposer d'une base de comparaison fixe, nous avons limité notre analyse à la bande spectrale correspondant aux fréquences les plus significatives pour la compréhension de la parole, c'est-à-dire les fréquences comprises entre 250 et 8000 Hz. Le choix de la limite basse à 250 Hz au lieu de 125 Hz est justifié par le fait que la position de la cochlée codant pour 125 Hz se situe au milieu du deuxième tour de cochlée (voir fig. 1.10, p. 25), et n'est donc la plupart du temps pas stimulée par les électrodes de l'implant cochléaire.

La détermination des fréquences de coupure se fait donc selon les étapes suivantes :

1. choix des fréquences extrêmales (dans nos études, 250 et 8000 Hz) ;
2. à partir de la carte tonotopique de Greenwood, calcul des positions extrêmales correspondantes sur la cochlée (respectivement 6.67 et 28.22 mm depuis l'apex) ;
3. en fonction du nombre d'électrodes à simuler, subdivision équidistante de la cochlée entre ces deux positions extrêmales ;
4. à partir de la carte tonotopique de Greenwood, calcul des fréquences de coupure correspondant à ces subdivisions équidistantes de la cochlée.

Décomposition spectrale

L'étape suivante consiste à se placer dans l'espace des fréquences de Fourier et à extraire de façon sélective les composantes spectrales dans chaque bande fréquentielle. Cette analyse fréquentielle se fait à l'aide de filtres passe-bande elliptiques IIR du sixième ordre, les filtres elliptiques présentant une sélectivité en fréquence bien plus fine que les filtres classiques de type Butterworth (voir *décomposition en bandes spectrales*, fig. 10.1, p. 220).

Uniformisation de l'énergie spectrale dans chaque canal fréquentiel

Une électrode donnée de l'implant cochléaire induit une stimulation électrique quasi-identique de tous les points de la cochlée en contact avec cette électrode. Par conséquent, le signal électrique émis par l'implant dans la bande spectrale associée à cette électrode est identique quelle que soit la fréquence considérée dans cette bande spectrale. Pour le formuler en d'autres termes, l'énergie portée par le signal physiologique résultant possède à chaque instant une répartition spectrale uniforme dans chaque bande de fréquence.

Dans le cadre d'une simulation acoustique de l'implant cochléaire, il est donc nécessaire de rendre compte de cette uniformité spectrale dans les différents canaux fréquentiels associés aux électrodes de l'implant. L'opération d'uniformisation doit avoir pour effet d'égaliser à chaque instant la répartition spectrale de l'énergie sur toute la bande, tout en assurant la préservation de l'énergie acoustique totale de la bande spectrale (voir *uniformisation de l'énergie acoustique dans chaque bande*, fig. 10.1, p. 220).

Afin de mener à bien cette opération d'uniformisation, il est indispensable d'obtenir dans un premier temps l'évolution de l'énergie acoustique totale de la bande spectrale en fonction

du temps. Cette évolution est fournie par l'enveloppe temporelle du signal de bande. Une fois connue cette information, une façon simple de redistribuer l'énergie uniformément sur toutes les fréquences de la bande spectrale est d'utiliser l'enveloppe temporelle pour moduler un bruit blanc. Comme le signal ainsi obtenu couvre toute l'étendue du spectre auditif, il sera nécessaire d'ajouter à la suite une dernière étape de traitement destinée à restreindre ce signal à la bande spectrale de travail. Il est intéressant de noter que cet ensemble d'opérations a pour effet de retirer les informations de structure fine de chaque signal de bande, tout en préservant les informations d'enveloppe temporelle (voir à ce sujet *« Rôle de l'enveloppe temporelle et de la structure fine »*, § 9.2.2, p. 194), de sorte que la procédure SSSN fournit un paradigme expérimental permettant d'étudier l'influence des indices d'enveloppe temporelle seuls, avec une exclusion des indices de structure fine.

En pratique, pour chaque bande spectrale, l'extraction de l'enveloppe temporelle du signal de bande est effectuée par un redressement de demi-onde (on rend positive la partie négative du signal) puis par un filtrage passe-bas à 500 Hz à l'aide d'un filtre elliptique IIR du troisième ordre (voir fig. 10.2, p. 223). L'enveloppe résultante est ensuite utilisée pour moduler un bruit blanc délivré par un générateur de nombres pseudo-aléatoire, ce qui aboutit à un signal possédant les mêmes variations temporelles que le signal original sur toutes les fréquences du spectre auditif (voir fig. 10.2, p. 223). Ce signal est finalement réajusté spectralement à l'aide du même filtre passe-bande utilisé pour la sélection des bandes spectrales (elliptique IIR du sixième ordre) afin d'éviter une superposition avec les autres bandes spectrales (voir fig. 10.2, p. 223).

Recomposition spectrale

Les bruits modulés par les enveloppes de chaque bande sont ensuite recombinés additivement (*recomposition en signal vocodé*, fig. 10.1, p. 220). Enfin, un réajustement du niveau sonore moyen est effectué pour assurer l'équivalence des intensités sonores entre le son original et le son filtré.

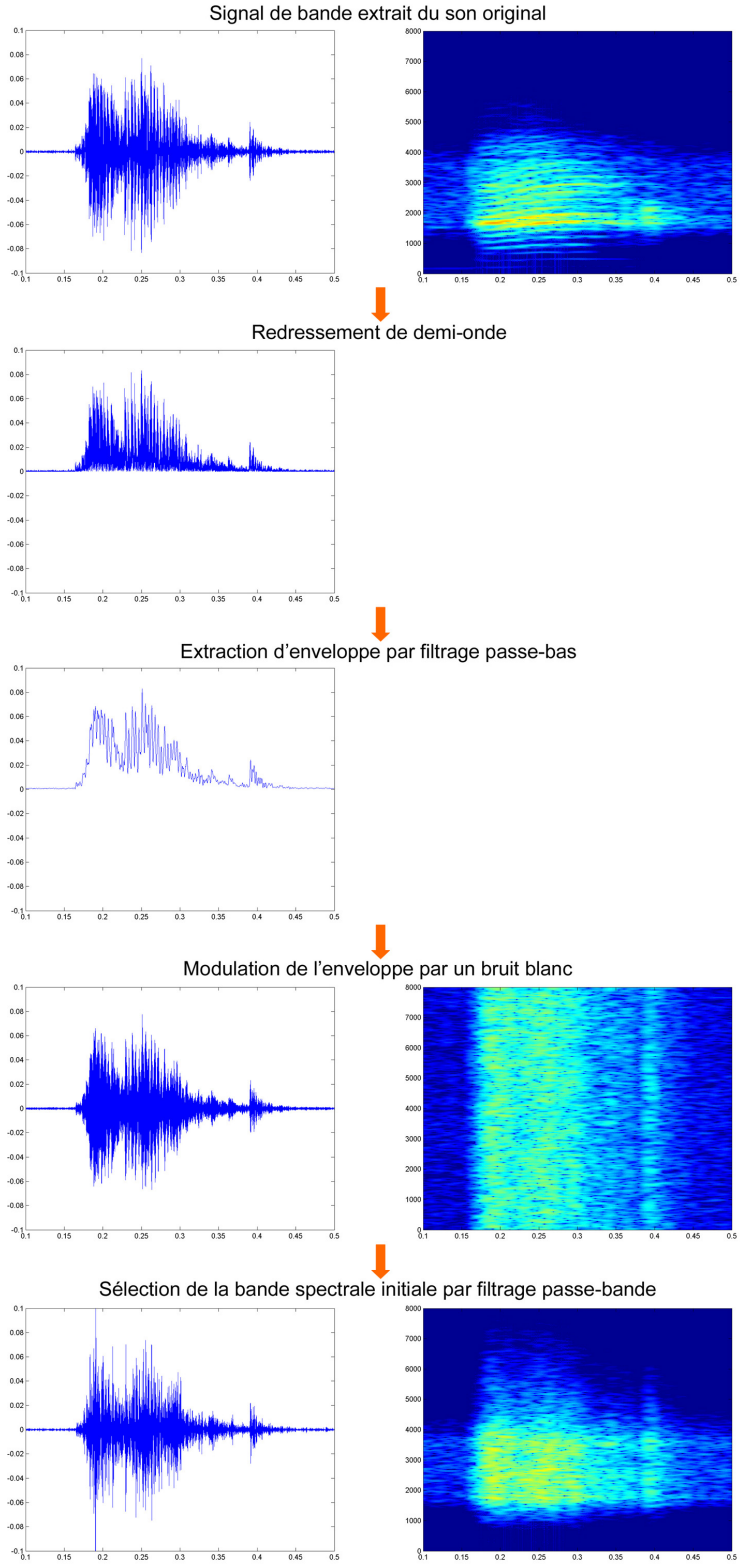


FIG. 10.2 – Etapes de transformation du signal acoustique lors de l'uniformisation de l'énergie spectrale dans chaque canal fréquentiel. Le signal sonore représenté ici est la première syllabe /ba/ du mot *balcon* utilisé dans la figure précédente. La bande de travail correspond à la bande 3 de la figure précédente.

10.3 Traitement phonétique des réponses comportementales

10.3.1 Analyse phonétique des réponses

Dans chacune des expériences comportementales faisant intervenir des mots dissyllabiques, toutes les réponses des sujets ont été transcrites dans un système phonétique simplifié. Les équivalences phonétiques utilisées sont décrites dans les tableaux 10.1 et 10.2.

I	/i/	l*i*t
E	/e/	d*é*
	/ɛ/	d*ai*s
A	/a/	t*a*
	/ɑ/	t*a*s
O	/ɔ/	s*o*rt
	/o/	s*o*t
OU	/u/	l*ou*p

U	/y/	l*u*
EU	/ø/	d*eu*x
	/œ/	p*eu*r
	/ə/	d*e*
AN	/ɑ̃/	d*en*t
ON	/ɔ̃/	d*on*
IN	/ɛ̃/	br*in*
	/œ̃/	br*un*

TAB. 10.1 – Système vocalique simplifié

Occlusives		
P	/p/	*p*ont
B	/b/	*b*on
M	/m/	*m*on
T	/t/	*t*emps
D	/d/	*d*ent
N	/n/	*n*eige
K	/k/	*c*lair
G	/g/	*g*ant

Constrictives		
F	/f/	*f*er
V	/v/	*v*erre
S	/s/	*s*aisir
Z	/z/	sai*s*ir
H	/ʃ/	*ch*amp
J	/ʒ/	*J*ean

Glides		
Y	/j/	c*i*el
W	/ɥ/	l*u*i
	/w/	L*ou*is

Liquides		
R	/ʁ/	*r*ond
L	/l/	*l*ong

TAB. 10.2 – Système consonantique simplifié

L'analyse phonétique des stimuli et des réponses est effectuée de la façon suivante :

1. on effectue une transcription phonétique dans le système simplifié décrit ci-avant ;
2. on décompose le mot en syllabes grâce aux phonèmes vocaliques (chacun n'est en effet impliqué que dans une seule syllabe) ;
3. dans chacune des syllabes, on recherche les phonèmes non vocaliques situés *avant* le noyau vocalique dans l'ordre suivant (du plus au moins proche) : 1° glide ; 2° liquide ; 3° occlusive ou constrictive ;
4. dans chacune des syllabes, on recherche les phonèmes non vocaliques situés *après* le noyau vocalique dans l'ordre suivant (du plus au moins proche) : 1° liquide ; 2° occlusive ou constrictive.

En utilisant les notations OC pour un phonème occlusif ou constrictif, L pour une liquide, G pour une glide et V pour un phonème vocalique, on obtient la structure générale suivante pour une syllabe : [OC][L][G]V[L][OC]. Comme ce mode de décomposition respecte la charpente phonétique naturelle des mots, la méthode proposée permet d'effectuer une comparaison phonème par phonème cohérente et systématique entre les réponses et les stimuli.

10.3.2 Matrices de confusions phonétiques

Toutes les réponses ayant subi une analyse phonétique ont été également représentées sous forme de matrices de confusions phonétiques. Ces matrices sont indexées horizontalement pour les stimuli et verticalement pour les percepts. Pour un stimulus donné (en abscisse), le pourcentage de réponses pour chaque percept possible (en ordonnée) est figuré par une couleur s'étendant du bleu (0% des réponses) au rouge (100% des réponses).

Comme les stimuli sont indexés dans le même ordre que les percepts dans les matrices de confusions, une répartition des réponses parfaitement diagonale indique une perception sans aucune confusion. Ainsi, plus la répartition des réponses est proche d'une disposition diagonale, moins la perception est confuse, et inversement. Enfin, une répartition des réponses en blocs diagonaux indique que les confusions ne se produisent qu'au sein de groupes particuliers de stimuli, chacun de ces groupes spécifiques correspondant à un bloc diagonal. Cette représentation en matrices de confusion fournit par conséquent une façon intéressante d'analyser les confusions perceptives rencontrées par les sujets.

Par commodité, nous avons choisi d'indexer les stimuli en donnant la priorité aux modes puis aux lieux d'articulations ; ainsi l'ordre d'indexation respecte la hiérachie suivante : occlusives bilabiales : /p/(P), /b/(B), /m/(M) ; occlusives dentales : /t/(T), /d/(D), /n/(N) ; occlusives vélaires : /k/(K), /g/(G) ; constrictives labiodentales : /f/(F), /v/(V) ; constrictives palatales : /ʃ/(H), /ʒ/(J) ; constrictives alvéolaires : /s/(S), /z/(Z) ; liquides/sonantes : ʀ/(R), /l/(L) ; glides : /w/ ou /ɥ/(W), /j/(Y).

10.3.3 Analyse de la transmission d'information phonétique

Modèle de transmission d'information

Chaque fois que l'étude s'y prêtait, nous avons effectué une analyse de transmission d'information basée sur le calcul de l'information phonétique préservée au cours du processus de la perception des stimuli. Le cadre théorique utilisé est le même que dans Miller et Nicely (1955). Nous en rappelons ici les grandes lignes.

On considère un protocole expérimental associant un stimulus x avec une réponse y. En fonction du protocole expérimental, le stimulus x pourra prendre n valeurs différentes x_1, x_2, …, x_n. En l'occurence, dans le cadre de nos études comportementales, ces différentes valeurs correspondent à différents phonèmes. De la même façon, la réponse y pourra prendre m valeurs différentes y_1, y_2, …, y_m ; ici également des phonèmes.

Chacun des n stimuli possibles a une probabilité p_i d'apparaître, fonction du protocole expérimental ; de même chacune des m réponses possibles intervient avec une probabilité p_j. Selon la théorie de l'information, la quantité d'information σ_x contenue dans le stimulus x (ou *entropie* de x) est définie comme l'espérance mathématique du logarithme de la probabilité :

$$\sigma_x = E(-\log p_i) = -\sum_{i=1}^{n} p_i \log p_i$$

De même, la quantité d'information σ_y contenue dans la réponse y est définie par :

$$\sigma_y = E(-\log p_j) = -\sum_{j=1}^{m} p_j \log p_j$$

Si on considère un logarithme en base 2, alors la quantité d'information ainsi calculée peut-être considérée comme le nombre de décisions binaires nécessaires pour définir complètement le

stimulus (ou la réponse), ou encore comme le nombre de bits d'information contenus dans le stimulus (ou la réponse).

On définit également la quantité d'information σ_{xy} associée aux couples stimulus-réponse possibles. Si p_{ij} désigne la probabilité d'occurence d'un couple $(x_i; y_j)$ donné, alors on a :

$$\sigma_{xy} = E(-\log p_{ij}) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} \log p_{ij}$$

Les quantités d'information σ_x, σ_y et σ_{xy} permettent alors de calculer une mesure de la covariance entre le stimulus x et la réponse y :

$$\mathrm{cov}(x, y) = \sigma_x + \sigma_y - \sigma_{xy} = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} \log \frac{p_i p_j}{p_{ij}}$$

La covariance $\mathrm{cov}(x, y)$ est également désignée comme la transmission d'information entre x et y exprimée en bits par stimulus. Cette quantité est positive et ne peut pas dépasser la quantité σ_x ; l'information recueillie en sortie du processus perceptif ne peut en effet pas être supérieure à l'information injectée à l'entrée. Si la covariance est égale à σ_x, cela signifie que toute l'information présente dans le stimulus a été préservée au cours du processus perceptif : les réponses sont alors parfaitement corrélées aux stimuli. Si la covariance est nulle, cela signifie que toute l'information présente dans le stimulus a été perdue au cours du processus perceptif : les réponses sont alors complètement indépendantes des stimuli. Afin de disposer d'une grandeur normalisée, on utilise usuellement la quantité σ_x^0, comprise entre 0 et 1, déduite de la covariance d'après la relation suivante :

$$\sigma_x^0 = \frac{\mathrm{cov}(x, y)}{\sigma_x}$$

Classification phonétique

Nous avons distingué plusieurs traits articulatoires caractéristiques au cours de notre analyse de transmission d'information. Concernant le lieu d'articulation, nous avons analysé la transmission de l'antériorité (phonèmes labiaux, labiodentaux et arrondis) et de la postériorité. Nous avons également analysé la transmission du voisement et de la nasalité, ainsi que la transmission de la constriction et de l'occlusion (voir la classification correpondante des consonnes, tab. 10.3, p. 226).

	P	**B**	**M**	**T**	**D**	**N**	**K**	**G**	**F**	**V**	**S**	**Z**	**H**	**J**	**Y**	**W**	**R**	**L**
Voisement	1	2	2	1	2	2	1	2	1	2	1	2	1	2	2	2	2	2
Nasalité	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
Mode	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3
Labialité	1	1	1	2	2	2	2	2	1	1	2	2	1	1	2	2	2	2

TAB. 10.3 – Traits articulatoires utilisés pour l'analyse de transmission d'information phonétique. Pour le voisement, on distingue les phonèmes non voisés (1) des voisés (2) ; pour la nasalité, on distingue les phonèmes non nasaux (1) des nasaux (2) ; pour le mode d'articulation, on distingue les occlusives (1) des constrictives (2) et des glides/liquides (3) ; enfin pour la labialité on distingue les phonèmes labiaux (1) des non-labiaux (2).

10.4 Neuroimagerie fonctionnelle

L'objectif de nos études neurofonctionnelles est d'étudier les conséquences de l'implantation cochléaire sur la réorganisation des réseaux corticaux impliqués dans la perception audiovisuelle de la parole. En raison de la présence d'un aimant dans le corps de l'implant cochléaire il existe des contre-indications médicales pour l'examen des patients implantés en imagerie fonctionnelle par résonance magnétique. Nous avons donc défini des protocoles de neuroimagerie en tomographie par émission de positrons, cette technique d'imagerie ne présentant aucune contre-indication particulière.

10.4.1 Déroulement d'un examen

Une séance de tomographie se déroule de la façon suivante. Le patient est allongé sur une couchette dans l'axe du tomographe, la tête soutenue confortablement et maintenue en place par un bandeau élastique. Les sons sont diffusés par deux haut-parleurs situés de part et d'autre de l'écran de stimulation. Une commande à deux boutons (souris d'ordinateur) est placée dans la main droite du patient pour lui permettre de répondre aux stimulations. Dans notre étude (cf. « *Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral* », ch. 13, p. 289), un écran de stimulation est installé dans l'axe du scanner, au-delà de la tête du patient, et les images affichées sur l'écran sont retransmises par un système de deux miroirs réglables orientés afin de permettre une vision confortable depuis la position allongée.

Les stimulations auditives, visuelles et audiovisuelles ainsi que les références motrices sont toutes structurées de la même façon. L'expérimentateur explique une nouvelle fois les instructions au patient, par oral et à l'aide d'un support écrit, en lui rappelant la signification associée à chacun des boutons de réponse. Puis une séquence vidéo sans son est présentée à l'écran et reprend à nouveau ces instructions. A la suite de cette séquence, après un décompte visuel de 10 secondes, la série de stimulations commence : chaque stimulus individuel est présenté pendant une durée moyenne de 1.5 s puis suivi d'un temps mort de 3.5 s sans son ni image, pendant lequel le patient peut appuyer sur le bouton de son choix pour indiquer sa réponse.

10.4.2 Dispositif expérimental

Les examens cliniques réalisés dans le cadre de nos études neuro-fonctionnelles ont eu lieu dans le centre de tomographie par émission de positrons du CHU Toulouse-Purpan. L'acquisition des images est effectuée par une caméra cylindrique EXACT.HR (CTI/Siemens) comprenant quatre anneaux de 72 blocs de détecteurs avec 64 cristaux par détecteur. Ce tomographe permet l'acquisition simultanée de 63 coupes transaxiales avec une résolution spatiale avant reconstruction de 2.4 mm en axial et 2.2 mm en transaxial. Les images sont reconstruites après acquisition par rétroprojection avec un filtrage de Hanning (fréquence de coupure à 0.4 cycle par projection) pour 63 coupes de 128×128 pixels. La résolution après reconstruction est de 4.5 mm dans le plan de l'image et de 4.1 mm en axial.

La voie d'administration du radiopharmaceutique nécessite la mise en place par une infirmière d'un cathéter veineux. De l'eau marquée à l'oxygène 15 (H_2O^{15}) dépourvue de toute propriété pharmacologique est injectée sous la responsabilité conjointe du radiopharmacien et du médecin nucléaire. La quantité globale de radioactivité correspondant à un examen reste inférieure à 5 mSv. Dans les limites des connaissances scientifiques actuelles il n'existe aucune pathologie radioinduite pour une telle dose de radioactivité.

Une injection intraveineuse en bolus de 6 mCi d'H_2O^{15} est réalisée 10 secondes avant le début des acquisitions puis pendant la durée de chaque acquisition. Pour notre étude (cf.

FIG. 10.3 – Centre de tomographie par émission de positrons du CHU Toulouse-Purpan

« *Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral* », ch. 13, p. 289), chaque séance de tomographie inclut 8 injections de radiopharmaceutique, ce qui permet de réaliser dans la même année les deux séances d'imagerie fonctionnelle TEP prévues pour le suivi longitudinal.

10.4.3 Prétraitement des images d'activation cérébrale

Toutes les images acquises pendant les tomographies sont ensuite analysées sous le logiciel SPM (statistical parametric mapping)†, piloté depuis un environnement Matlab. Avant l'analyse statistique, les images subissent :

1. un réalignement (en translation et en rotation) afin de compenser les légers déplacements des sujets au cours de la séance de tomographie ;
2. une renormalisation anatomique par transformation affine sur une image cérébrale de référence, afin de pouvoir disposer d'un repère anatomique identique pour tous les sujets. L'image de référence utilisée est celle du Montreal Neurological Institute (voir « *Normalisation de l'anatomie corticale* », sect. A.2, p. 347) ;
3. un lissage des images par une gaussienne tridimensionnelle (un écart-type de 8 mm est utilisé dans les trois dimensions). Ce lissage permet : 1° de se départir des variabilités anatomiques interindividuelles qui n'ont pas été compensées par la renormalisation ; 2° de se rapprocher des conditions théoriques qui permettent une analyse statistique optimale.

Bibliographie

Blamey, P. J., Dooley, G. J., Parisi, E. S. et Clark, G. M. (1996). Pitch comparisons of acoustically and electrically evoked auditory sensations. *Hear. Res.*, 99(1-2):139–150.

Boex, C., Baud, L., Cosendai, G., Sigrist, A., Kos, M. I. et Pelizzone, M. (2006). Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J. Assoc. Res. Otolaryngol.*, 7(2):110–124.

Dorman, M. F., Spahr, T., Gifford, R., Loiselle, L., McKarns, S., Holden, T., Skinner, M. et Finley, C. (2007). An Electric Frequency-to-place Map for a Cochlear Implant Patient with Hearing in the Nonimplanted Ear. *J. Assoc. Res. Otolaryngol.*, 8(2):234–240.

Dudley, H. (1939a). Remaking Speech. *J. Acoust. Soc. Am.*, 11(2):169–177.

Dudley, H. (1939b). The vocoder. *In Bell Labs Record*, pages 122–126.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species - 29 years later. *J. Acoust. Soc. Am.*, 87:2592–2605.

James, C., Blamey, P. J., Shallop, J. K., Incerti, P. V. et Nicholas, A. M. (2001). Contralateral masking in cochlear implant users with residual hearing in the non-implanted ear. *Audiol. Neurootol.*, 6(2):87–97.

Miller, G. A. et Nicely, P. (1955). An analysis of perceptual confusions among some english consonants. *J. Acoust. Soc. Am.*, 27(2):338–352.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonsky, J. et Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270:303–304.

Stakhovskaya, O., Sridhar, D., Bonham, B. H. et Leake, P. A. (2007). Frequency map for the human cochlear spiral ganglion : implications for cochlear implants. *J. Assoc. Res. Otolaryngol.*, 8(2):220–233.

Chapitre 11

Etudes comportementales chez le sujet sain

Sommaire

11.1 Lecture labiale chez le sujet sain 232
11.1.1 Matériel et méthodes 232
11.1.2 Résultats 233
11.1.3 Discussion 233
11.2 Sujet sain : parole masquée par du bruit 235
11.2.1 Matériel et méthodes 235
11.2.2 Résultats 236
11.2.3 Discussion 242
11.3 Sujet sain : parole traitée par l'implant 242
11.3.1 Matériel et méthodes 242
11.3.2 Résultats 244
11.3.3 Discussion 248
11.4 Entraînement perceptif avec une simulation d'implant 249
11.4.1 Matériel et méthodes 250
11.4.2 Résultats 251
11.4.3 Discussion 253
Bibliographie 255

Introduction

Afin d'explorer les mécanismes d'intégration audiovisuelle de la parole utilisés par les patients implantés cochléaires, nous avons défini plusieurs protocoles comportementaux originaux destinés à évaluer la compréhension de la parole des sujets contrôles normo-entendants dans des conditions perceptives se rapprochant de celles vécues par les patients sourds implantés. Ces protocoles fournissent une base de comparaison apportant un éclairage particulier sur la perception audiovisuelle de la parole chez les patients implantés.

Ainsi, dans une première étude (voir *« Lecture labiale chez le sujet sain »*, sect. 11.1, p. 232), nous évaluons les performances en lecture labiale des contrôles normo-entendants

(n_{NH}= 42). Afin de se rapprocher de la dégradation du signal auditif par l'implant cochléaire, nous les testons ensuite dans deux conditions de dégradation de l'audition, d'une part en appliquant un bruit blanc masquant aux signaux de parole (n_{NH}= 80 ; voir *« Sujet sain : parole masquée par du bruit »*, sect. 11.2, p. 235), d'autre part en traitant ces signaux au travers d'une simulation acoustique d'implant cochléaire (n_{NH}= 41 ; voir *« Sujet sain : parole traitée par l'implant »*, sect. 11.3, p. 242). Enfin, dans une dernière étude, nous évaluons l'effet d'un entraînement auditif ou audiovisuel à moyen terme chez des contrôles normo-entendants en les soumettant à une simulation du traitement des signaux de parole par l'implant cochléaire (n_{NH}= 14 ; voir *« Entraînement perceptif avec une simulation d'implant »*, sect. 11.4, p. 249).

L'utilisation de simulations acoustiques d'implant cochléaire en condition audiovisuelle constitue un paradigme expérimental original qui n'a à notre connaissance pas encore été employé dans d'autres programmes de recherche.

11.1 Lecture labiale chez le sujet sain

11.1.1 Matériel et méthodes

Sujets

Cette investigation a été effectuée dans le centre Cerveau et Cognition UMR CNRS 5549 à la faculté de médecine de Toulouse-Rangueil. Nous avons réalisé une étude prospective sur 42 sujets adultes ne présentant aucune déficience auditive. Le groupe de sujets est composé de 20 femmes et 22 hommes entre 22 et 58 ans, avec une moyenne d'âge de 33.3 ans (± 10.5 ans). Tous les sujets étaient entièrement naïfs vis-à-vis du protocole expérimental, aucun d'entre eux n'ayant jamais été exposé auparavant aux stimuli utilisés.

Stimuli

Les stimuli sont tous des mots dissyllabiques faisant partie des listes normalisées de Fournier. Ces mots ont été répartis en trois listes de 20 mots chacune. Ces listes ont été équilibrées en fréquence d'utilisation lexicale, comptabilisée en nombre d'occurences par million de mots d'après la base de données Brulex (Content *et al.*, 1990), en structure phonétique (répartition voyelles antérieures/voyelles postérieures) et en structure syllabique (CV/CVC/CCV). L'enregistrement des prononciations de chaque mot s'est déroulé suivant les modalités décrites dans *« Génération des stimuli »*, § 10.2.1, p. 218.

Procédure

Les stimuli ont été présentés sur un ordinateur portable avec un moniteur de dimensions 20.7×33.1 cm. L'image était présentée au centre de l'écran avec une dimension fixe de 20.7×27.6 cm, le reste de l'écran demeurant noir. Les sujets étaient assis confortablement en face de l'écran, à une distance de lecture normale d'environ 65 cm, ce qui correspond à un diamètre angulaire d'environ 30°. Le son était délivré à travers un casque d'écoute avec une faible déformation spectrale, de telle sorte que le niveau sonore moyen atteigne 65 dB.

Les participants ont tous été testés individuellement en condition de lecture labiale sans le son. Ils ont reçu pour seules instructions de regarder l'image pendant toute la durée de présentation de chaque stimulus, puis de répéter à voix haute ce qu'ils avaient perçu. L'expérimentateur

était assis dans la pièce pendant toute la durée de l'expérience pour confirmer que les sujets regardaient bien l'écran et pour consigner par écrit leurs réponses verbales.

Analyse phonétique des confusions

Les réponses des sujets ont subi une analyse phonétique complète, suivant les méthodes décrites dans *« Analyse phonétique des réponses »*, § 10.3.1, p. 224. Les résultats de cette analyse ont été représentés sous forme de matrices de confusions phonétiques, dont le principe est détaillé dans *« Matrices de confusions phonétiques »*, § 10.3.2, p. 225. La stimulation en condition visuelle incluait 60 mots disyllabiques contenant 143 phonèmes consonantiques et 120 phonèmes vocaliques, soit 263 phonèmes au total, ce qui correspond à une moyenne de 4.4 phonèmes par mot. Nous avons également effectué une analyse de transmission d'information à partir de l'analyse phonétique des réponses des sujets ; les principes de base de cette analyse sont résumés dans *« Analyse de la transmission d'information phonétique »*, § 10.3.3, p. 225.

11.1.2 Résultats

Performances globales

Les performances de reconnaissance des mots en lecture labiale observées dans notre groupe de 42 sujets atteignent un score de 9.4 ± 7.1%. L'analyse phonétique de la reconnaissance des phonèmes individuels indique des scores de 32.8 ± 11.0% pour les voyelles et 24.5 ± 9.3% pour les consonnes, les voyelles étant significativement mieux perçues que les consonnes ($p < 0.0001$; Wilcoxon en mesures appariées).

Différences hommes/femmes

Lorsque l'on effectue une analyse distincte pour le sous-groupe de 22 hommes et le sous-groupe de 20 femmes, on trouve des scores globaux de reconnaissance de mots respectivement égaux à 6.6 ± 2.9% (hommes) et 12.3 ± 5.9% (femmes), les femmes étant plus performantes que les hommes ($p < 0.0008$; Mann-Whitney).

Cette différence hommes/femmes reste valable au niveau des consonnes, avec des scores de 19.3 ± 6.5% pour les hommes et 30.4 ± 8.5% pour les femmes ($p < 0.0001$; Mann-Whitney), ainsi que pour les voyelles, avec des scores de 27.8 ± 9.9% pour les hommes et 38.2 ± 9.6% pour les femmes ($p < 0.003$; Mann-Whitney).

Par ailleurs, dans les deux sous-groupes, les voyelles sont toujours significativement mieux perçues que les consonnes ($p < 0.0001$; Wilcoxon en mesures appariées).

11.1.3 Discussion

Notre étude confirme le faible niveau de performance de lecture labiale atteint par les sujets normo-entendants (Risberg et Lubker, 1978; Goh *et al.*, 2001). Elle indique également qu'en condition de parole visuelle, les voyelles tendent à être mieux reconnues que les consonnes. Un premier élément d'explication possible de cette différence est la durée de prononciation des phonèmes, très rapide pour les consonnes et nettement plus longue pour les voyelles. Ainsi il se peut que la reconnaissance visuelle des voyelles soit facilitée par leur plus importante durée de prononciation, en comparaison avec la brièveté des consonnes. Cette différence voyelles/consonnes peut également être conditionnée par la plus grande diversité du système phonétique français en consonnes (17 consonnes et 3 glides) qu'en voyelles (16 voyelles ; voir *« Notation phonétique »*, § 2.2.2, p. 48), ce qui peut contribuer à augmenter la possibilité de confusions pour les consonnes.

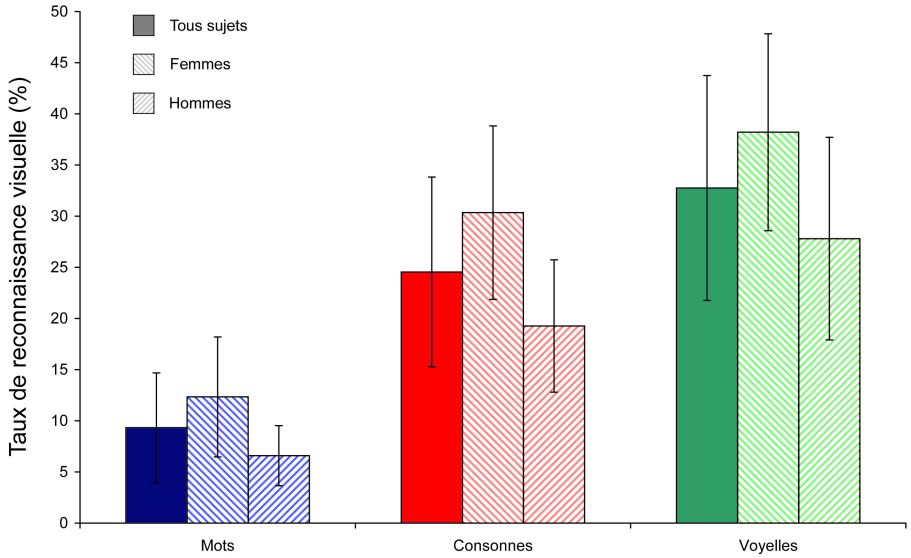


FIG. 11.1 – Performances de reconnaissance observées dans un groupe de 42 sujets normo-entendants (22 hommes et 20 femmes) exposés à de la parole visuelle.

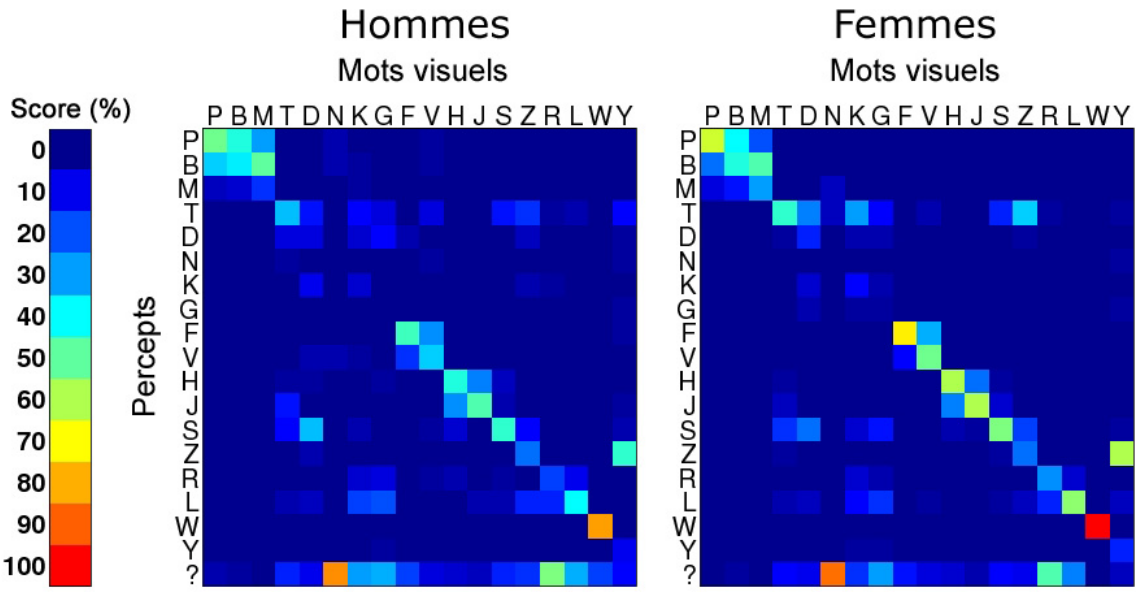


FIG. 11.2 – Analyse phonétique des confusions de consonnes dans des mots disyllabiques présentés en vision seule observées dans un groupe de 22 hommes et 20 femmes.

Cette étude suggère également des différences entre les mécanismes de traitement de la parole utilisés par les femmes et les hommes, les femmes parvenant à un niveau de reconnaissance nettement supérieur à celui des hommes, aussi bien au niveau des mots entiers que des phonèmes consonantiques et vocaliques qui les composent. Ces résultats sont cohérents avec un ensemble d'études indiquant que les femmes obtiennent des performances de lecture labiale significativement plus importantes que les hommes (Johnson et Leslie, 1988; Watson *et al.*, 1996). Il sont également à mettre en parallèle avec certaines études démontrant une plus grande influence de la

composante visuelle de la parole chez les femmes lors de la perception de stimuli de parole audiovisuelle incongruente de type McGurk (Aloufy *et al.*, 1996; Irwin *et al.*, 2006) (voir *« Différence d'intégration audiovisuelle entre hommes et femmes »*, § 3.6.4, p. 71).

Nous développons une discussion plus détaillée de cet ensemble de résultats en comparant les sujets normo-entendants aux patients implantés dans *« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258 puis en analysant de façon approfondie les différences hommes/femmes dans *« Lecture labiale chez le sujet implanté »*, sect. 12.3, p. 268.

11.2 Sujet sain : parole masquée par du bruit

11.2.1 Matériel et méthodes

Sujets

Cette investigation a eu lieu dans le centre Cerveau et Cognition UMR CNRS 5549 à la faculté de médecine de Toulouse-Rangueil. Nous avons réalisé une étude prospective sur 80 sujets adultes ne présentant aucune déficience auditive. Le groupe de sujets est composé de 46 femmes et 34 hommes entre 17 et 61 ans, avec une moyenne d'âge de 24.8 ans ($\pm$ 8.7 ans). Tous les sujets étaient entièrement naïfs vis-à-vis du protocole expérimental, aucun d'entre eux n'ayant jamais été exposé auparavant aux stimuli utilisés.

Stimuli

Sélection des stimuli : Les stimuli sont tous des mots dissyllabiques faisant partie des listes normalisées de Fournier. Ces mots ont été répartis en cinq listes de 20 mots chacune. Ces listes ont été équilibrées en fréquence d'utilisation lexicale, comptabilisée en nombre d'occurences par million de mots d'après la base de données Brulex (Content *et al.*, 1990), en structure phonétique (répartition voyelles antérieures/voyelles postérieures) et en structure syllabique (CV/CVC/CCV). L'enregistrement des prononciations de chaque mot s'est déroulé suivant les modalités décrites dans *« Génération des stimuli »*, § 10.2.1, p. 218.

Génération des stimuli : Les sons ont été travaillés sous Matlab 6.5 sur un processeur Pentium 4 2.4 GHz. Le masque utilisé dans cette expérience est un bruit blanc engendré par un générateur de nombres pseudo-aléatoire et modulé temporellement par des lobes sinusoïdaux monopériodiques de 20 ms, l'atténuation au centre des lobes étant de 40%. Un taux de 300 modulations par seconde a été utilisé, chacune des modulations étant appliquée dans une fenêtre temporelle de 20 ms sélectionnée aléatoirement.

Les stimuli ont été montés ensemble sous forme de films vidéo au format MPEG2 avec Adobe Premiere Pro 7.0 sur un processeur Pentium 4 2.4 GHz. Trois conditions ont été testées : une condition auditive seule (A), pour laquelle seul le son de l'enregistrement original était présenté, l'écran restant totalement noir ; une condition auditive bruitée (Ab), pour laquelle le son de l'enregistrement original était masqué par du bruit blanc aléatoire ; puis une condition audiovisuelle bruitée (AVb) pour laquelle l'image de l'enregistrement original était présentée conjointement avec le son masqué correspondant.

Ordre de présentation : chaque sujet a été exposé aux cinq listes de vingt mots, une liste étant utilisée par test. Le premier test consistait en un test de compréhension auditive sans dégradation de l'audition, afin de contrôler la normalité de la compréhension. Les quatre autres présentations incluaient deux tests auditifs en condition bruitée et deux tests audiovisuels en condition bruitée, tous mélangés dans un ordre aléatoire.

Procédure

Les stimuli ont été présentés sur un ordinateur portable avec un moniteur de dimensions 20.7×33.1 cm. L'image était présentée au centre de l'écran avec une dimension fixe de 20.7×27.6 cm, le reste de l'écran demeurant noir. Les sujets étaient assis confortablement en face de l'écran, à une distance de lecture normale d'environ 65 cm, ce qui correspond à un diamètre angulaire d'environ 30°. Le son était délivré à travers un casque d'écoute avec une faible déformation spectrale, de telle sorte que le niveau sonore moyen atteigne 65 dB.

Les participants ont tous été testés individuellement. Ils ont reçu pour seules instructions de regarder l'image et d'écouter le son pendant toute la durée de présentation de chaque stimulus, puis de répéter à voix haute ce qu'ils avaient perçu. L'expérimentateur était assis dans la pièce pendant toute la durée de l'expérience pour confirmer que les sujets regardaient bien l'écran et pour consigner par écrit leurs réponses verbales.

Analyse phonétique des confusions

Les réponses des sujets ont subi une analyse phonétique complète, suivant les méthodes décrites dans *« Analyse phonétique des réponses »*, § 10.3.1, p. 224. Les résultats de cette analyse ont été représentés sous forme de matrices de confusions phonétiques, dont le principe est détaillé dans *« Matrices de confusions phonétiques »*, § 10.3.2, p. 225. La condition auditive incluait 40 mots disyllabiques contenant 95 phonèmes consonantiques et 80 phonèmes vocaliques, soit 175 phonèmes au total, ce qui correspond à une moyenne de 4.4 phonèmes par mot. La condition audiovisuelle incluait 40 mots disyllabiques contenant 98 phonèmes consonantiques et 80 phonèmes vocaliques, soit 178 phonèmes au total, ce qui correspond à une moyenne de 4.5 phonèmes par mot. Nous avons également effectué une analyse de transmission d'information à partir de l'analyse phonétique des réponses des sujets ; les principes de base de cette analyse sont résumés dans *« Analyse de la transmission d'information phonétique »*, § 10.3.3, p. 225.

11.2.2 Résultats

L'analyse des performances globales des sujets met en évidence une amélioration de la perception de la parole en condition audiovisuelle en comparaison avec la condition auditive seule, cette amélioration se produisant quel que soit le niveau de dégradation de la parole, pour des rapports signal/bruit (SNR) variant de 0 à -25 dB. Les niveaux d'identification observés pour les voyelles et les consonnes sont plus élevés que pour les mots, cette dernière façon d'évaluer les performances perceptives étant en effet moins permissive. De façon intéressante, les voyelles sont systématiquement mieux reconnues que les consonnes, quel que soit le niveau de dégradation de la parole.

On observe l'apparition d'un maximum du gain audiovisuel pour des SNR intermédiaires proches de -15 dB ; au-delà de cette zone le gain audiovisuel est d'autant plus faible que le SNR est faible (et donc que la parole est dégradée), tandis qu'en deçà de cette zone, le gain audiovisuel est d'autant plus faible que le SNR est élevé (et donc que la parole est préservée). Il est intéressant de remarquer un phénomène d'inversion des gains audiovisuels associés aux mots entiers et aux consonnes/voyelles, cette inversion se produisant également dans la zone des SNR voisins de -15 dB : pour des SNR plus élevés (parole mieux préservée), les mots bénéficient d'un meilleur gain audiovisuel que les phonèmes pris individuellement, tandis que pour des SNR plus faibles (parole plus dégradée), ce sont les voyelles et les consonnes individuelles qui bénéficient d'un gain audiovisuel plus élevé que les mots.

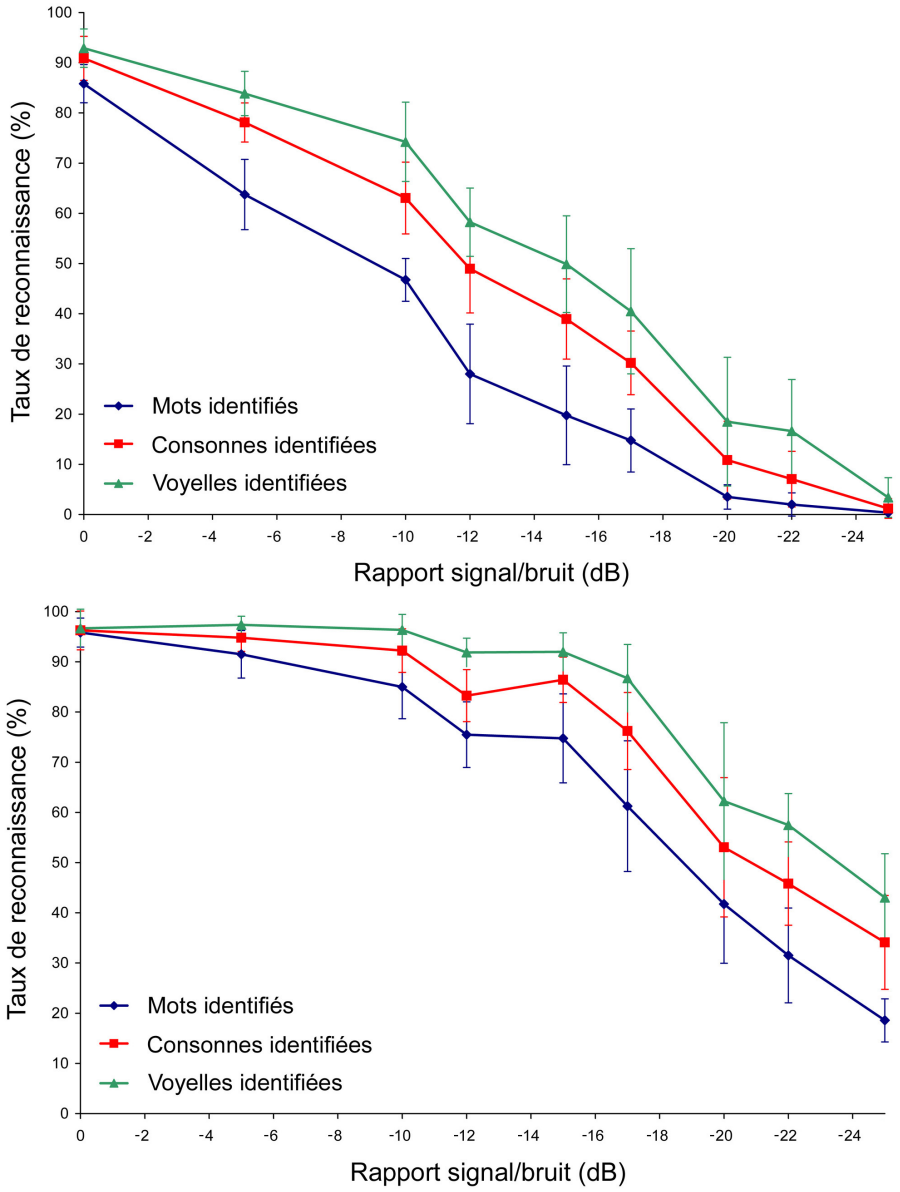


FIG. 11.3 – Performances de reconnaissance observées dans un groupe de 80 sujets normo-entendants exposés à de la parole auditive (en haut) et audiovisuelle (en bas) dégradée par un masque de bruit blanc.

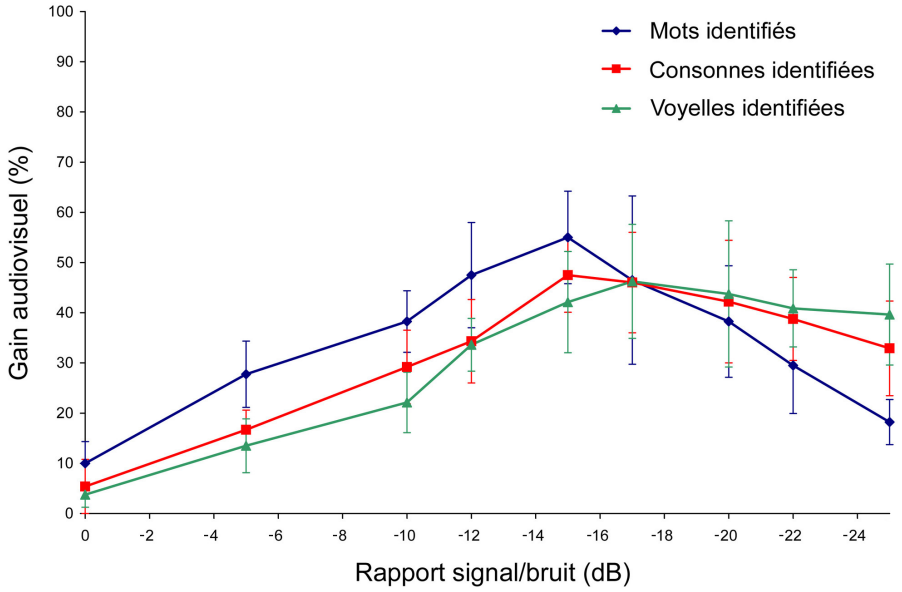


FIG. 11.4 – Gain audiovisuel observé dans un groupe de 80 sujets normo-entendants exposés à de la parole dégradée par un masque de bruit blanc.

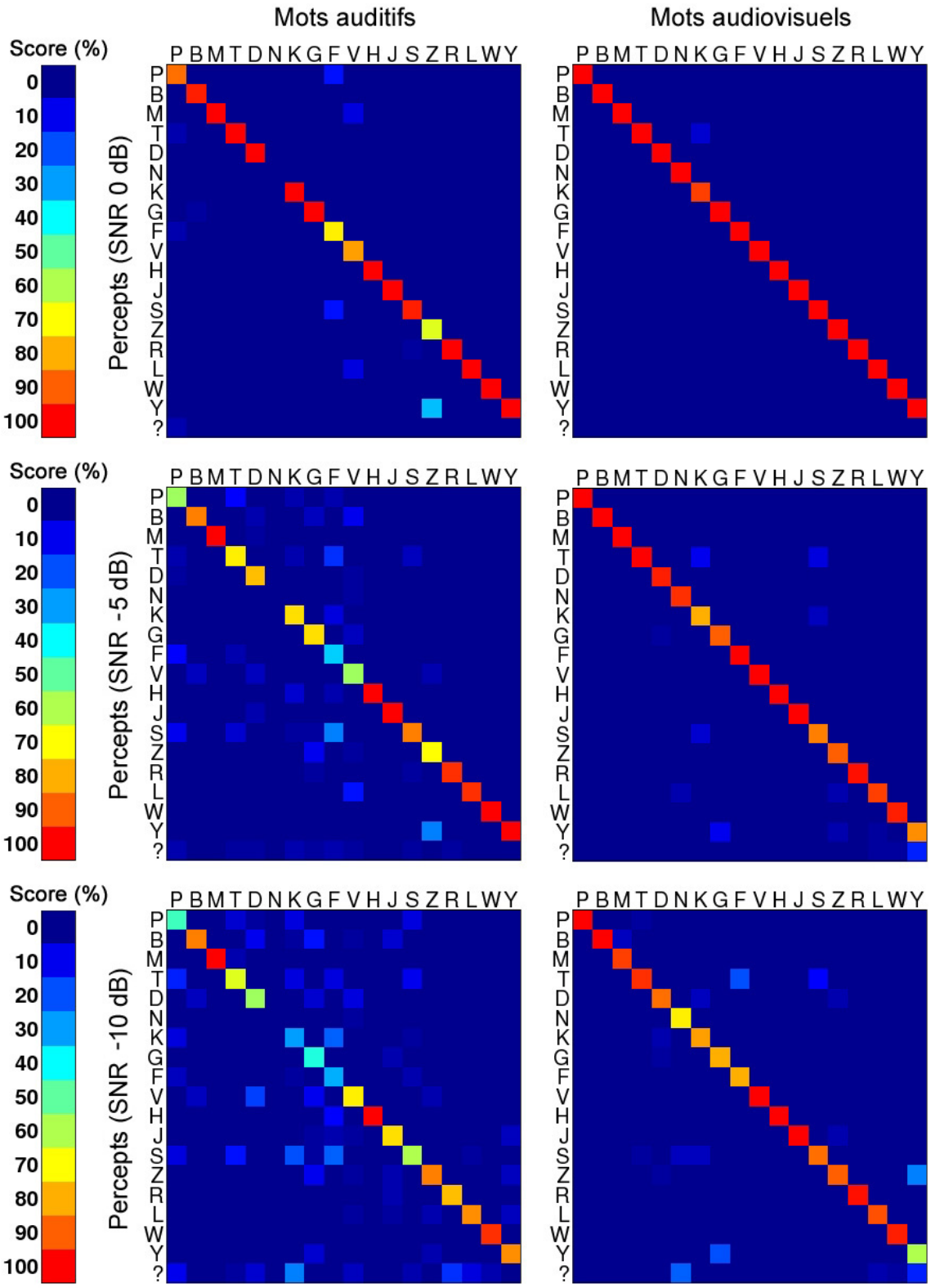


FIG. 11.5 – Analyse phonétique des confusions de consonnes réalisées par les sujets normaux sur des mots disyllabiques présentés en conditions auditive et audiovisuelle, avec un masquage par du bruit blanc. Rapport signal/bruit : 0 (n = 3), -5 dB (n = 10) et -10 dB (n = 10).

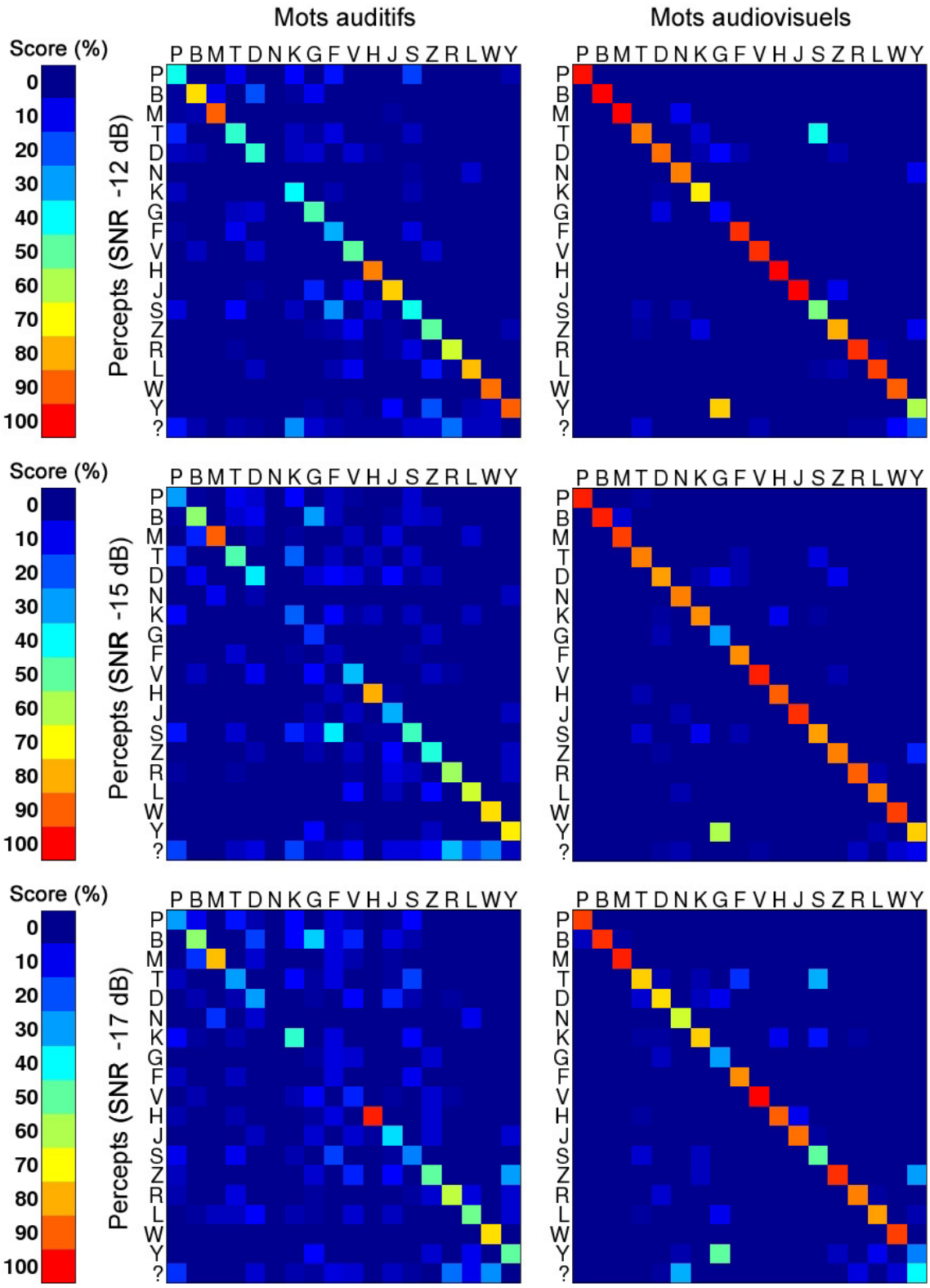


FIG. 11.6 – Analyse phonétique des confusions de consonnes réalisées par les sujets normaux sur des mots disyllabiques présentés en conditions auditive et audiovisuelle, avec un masquage par du bruit blanc. Rapport signal/bruit : -12 (n = 10), -15 dB (n = 10) et -17 dB (n = 10).

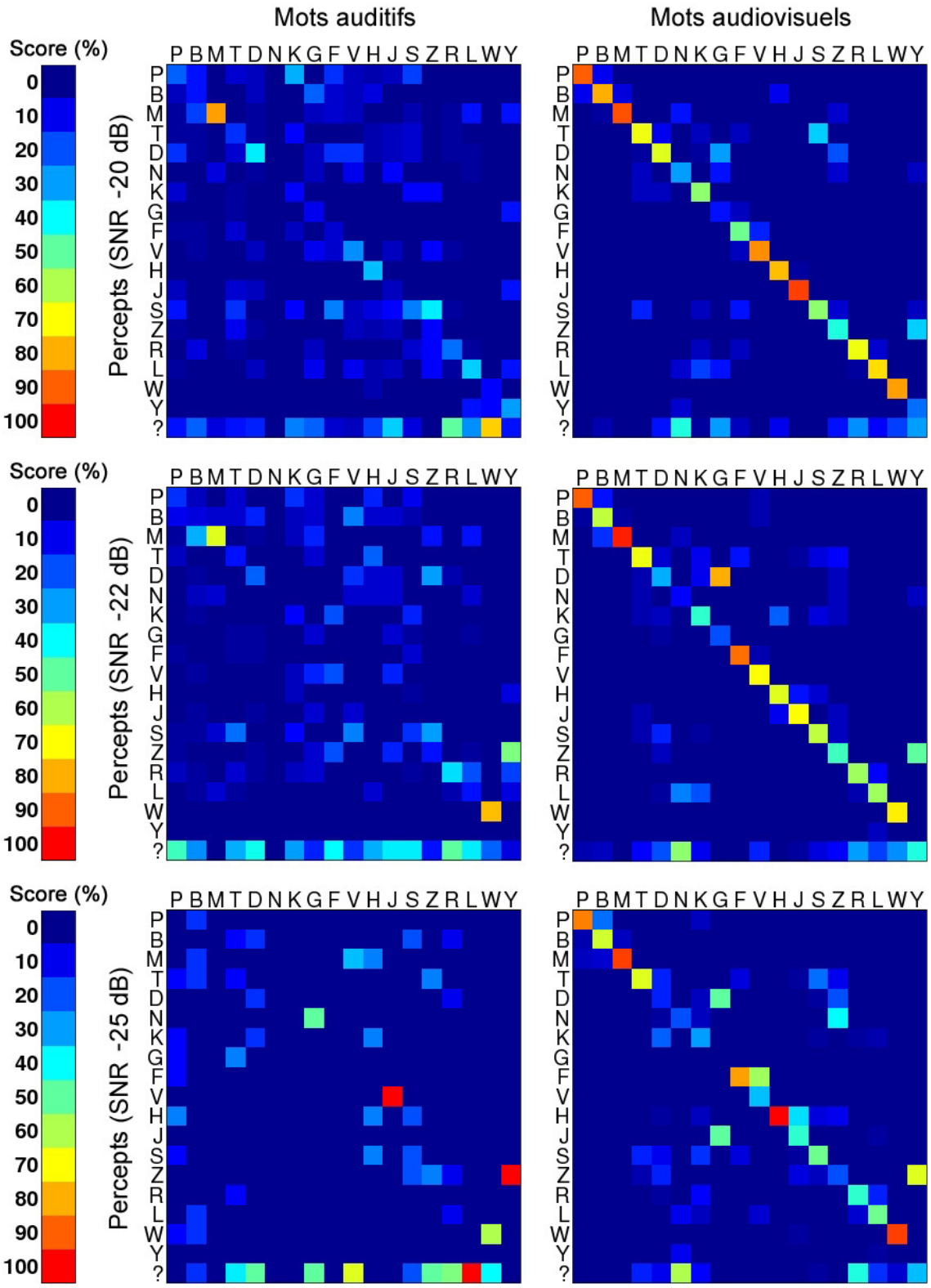


FIG. 11.7 – Analyse phonétique des confusions de consonnes réalisées par les sujets normaux sur des mots disyllabiques présentés en conditions auditive et audiovisuelle, avec un masquage par du bruit blanc. Rapport signal/bruit : -20 (n = 10), -22 dB (n = 10) et -25 dB (n = 7).

11.2.3 Discussion

Cette étude indique des résultats comportementaux cohérents avec les études portant sur la perception audiovisuelle de la parole en condition dégradée, qui démontrent le bénéfice apporté par une présentation audiovisuelle (Sumby et Pollack, 1954; Miller et Nicely, 1955; Neely, 1956; Erber, 1969, 1975; Binnie *et al.*, 1974; Summerfield, 1979; MacLeod et Summerfield, 1987; Grant et Braida, 1991; Benoit *et al.*, 1994) (voir *« Améliorations perceptives en condition dégradée »*, § 3.2.2, p. 61).

Ici la présentation en condition audiovisuelle permet une amélioration perceptive allant de 10% à plus de 50% suivant les rapports signal/bruit utilisés. Nous retrouvons la présence d'un gain audiovisuel maximal pour des rapports signal/bruit voisins de -15 dB, ce qui est cohérent avec la valeur de -12 dB indiquée dans la littérature (Ross *et al.*, 2007) (voir également *« Améliorations perceptives en condition dégradée »*, § 3.2.2, p. 61).

Nous discutons de façon détaillée de l'apport de ces données sur la compréhension des mécanismes d'intégration audiovisuelle de la parole dans *« Intégration audiovisuelle chez le sujet implanté »*, sect. 12.2, p. 262.

11.3 Sujet sain : parole traitée par l'implant

11.3.1 Matériel et méthodes

Sujets

Cette investigation a été effectuée dans le centre Cerveau et Cognition UMR CNRS 5549 à la faculté de médecine de Toulouse-Rangueil. Nous avons réalisé une étude prospective sur 41 sujets adultes ne présentant aucune déficience auditive. Le groupe de sujets est composé de 24 femmes et 17 hommes entre 17 et 56 ans, avec une moyenne d'âge de 23.7 ans (± 7.4 ans). Tous les sujets étaient entièrement naïfs vis-à-vis du protocole expérimental, aucun d'entre eux n'ayant jamais été exposé auparavant aux stimuli utilisés.

Stimuli

Sélection des stimuli : Les stimuli sont tous des mots dissyllabiques faisant partie des listes normalisées de Fournier. Ces mots ont été répartis en cinq listes de 20 mots chacune. Ces listes ont été équilibrées en fréquence d'utilisation lexicale, comptabilisée en nombre d'occurences par million de mots d'après la base de données Brulex (Content *et al.*, 1990), en structure phonétique (répartition voyelles antérieures/voyelles postérieures) et en structure syllabique (CV/CVC/CCV). L'enregistrement des prononciations de chaque mot s'est déroulé suivant les modalités décrites dans *« Génération des stimuli »*, § 10.2.1, p. 218.

Génération des stimuli : Les sons ont été travaillés sous Matlab 6.5 sur un processeur Pentium 4 2.4 GHz suivant une procédure SSSN (speech spectrum shaped noise). Ils ont été analysés dans différentes bandes spectrales à l'aide de filtres elliptiques IIR du sixième ordre suivant la carte tonotopique de Greenwood (Greenwood, 1990) ; seules les fréquences comprises entre 250 et 8000 Hz ont été retenues, avec un découpage variant de 2 à 16 bandes. Les fréquences correspondantes étaient : 250, 1676 et 8000 Hz pour la condition à deux bandes ; 250, 709, 1676, 3713 et 8000 Hz pour la condition à quatre bandes ; 250, 437, 709, 1104, 1676, 2507, 3713, 5462 et 8000 Hz pour la condition à huit bandes ; 250, 335, 437, 561, 709, 888, 1104, 1363, 1676, 2053, 2507, 3054, 3713, 4506, 5462, 6613 et 8000 Hz pour la condition à seize bandes.

Les opérations de traitement de signal utilisées pour la réalisation de ces stimuli sont décrites dans « *Simulation acoustique de l'implant cochléaire* », sect. 10.2.2, p. 218. On pourra consulter « *Influence de la profondeur d'implantation* », § 9.2.3, p. 196 pour plus de détails sur la procédure SSSN et pour des élements de discussion sur la validité d'utilisation de la carte tonotopique de Greenwood dans le cadre d'une simulation acoustique d'implant cochléaire.

Les stimuli ont été montés ensemble sous forme de films vidéo au format MPEG2 avec Adobe Premiere Pro 7.0 sur un processeur Pentium 4 2.4 GHz. Trois conditions ont été testées : une condition auditive seule (A), pour laquelle seul le son de l'enregistrement original était présenté, l'écran restant totalement noir ; une condition auditive filtrée (A-SSN) pour laquelle le son était filtré suivant la procédure décrite précédemment ; puis une condition audiovisuelle filtrée (AV-SSN) pour laquelle l'image de l'enregistrement original était présentée conjointement avec le son filtré correspondant.

Ordre de présentation : chaque sujet a été exposé aux cinq listes de vingt mots, une liste étant utilisée par test. Le premier test consistait en un test de compréhension auditive sans dégradation de l'audition, afin de contrôler la normalité de la compréhension. Les quatre autres présentations incluaient deux tests auditifs en condition filtrée et deux tests audiovisuels en condition filtrée, tous mélangés dans un ordre aléatoire.

Procédure

Les stimuli ont été présentés sur un ordinateur portable avec un moniteur de dimensions 20.7×33.1 cm. L'image était présentée au centre de l'écran avec une dimension fixe de 20.7×27.6 cm, le reste de l'écran demeurant noir. Les sujets étaient assis confortablement en face de l'écran, à une distance de lecture normale d'environ 65 cm, ce qui correspond à un diamètre angulaire d'environ 30°. Le son était délivré à travers un casque d'écoute avec une faible déformation spectrale, de telle sorte que le niveau sonore moyen atteigne 65 dB.

Les participants ont tous été testés individuellement. Ils ont reçu pour seules instructions de regarder l'image et d'écouter le son pendant toute la durée de présentation de chaque stimulus, puis de répéter à voix haute ce qu'ils avaient perçu. L'expérimentateur était assis dans la pièce pendant toute la durée de l'expérience pour confirmer que les sujets regardaient bien l'écran et pour consigner par écrit leurs réponses verbales.

Analyse phonétique des confusions

Les réponses des sujets ont subi une analyse phonétique complète, suivant les méthodes décrites dans « *Analyse phonétique des réponses* », § 10.3.1, p. 224. Les résultats de cette analyse ont été représentés sous forme de matrices de confusions phonétiques, dont le principe est détaillé dans « *Matrices de confusions phonétiques* », § 10.3.2, p. 225. La condition auditive incluait 40 mots disyllabiques contenant 95 phonèmes consonantiques et 80 phonèmes vocaliques, soit 175 phonèmes au total, ce qui correspond à une moyenne de 4.4 phonèmes par mot. La condition audiovisuelle incluait 40 mots disyllabiques contenant 98 phonèmes consonantiques et 80 phonèmes vocaliques, soit 178 phonèmes au total, ce qui correspond à une moyenne de 4.5 phonèmes par mot. Nous avons également effectué une analyse de transmission d'information à partir de l'analyse phonétique des réponses des sujets ; les principes de base de cette analyse sont résumés dans « *Analyse de la transmission d'information phonétique* », § 10.3.3, p. 225.

11.3.2 Résultats

L'analyse des performances globales des sujets met en évidence une amélioration de la perception de la parole en condition audiovisuelle en comparaison avec la condition auditive seule, cette amélioration se produisant quel que soit le niveau de dégradation de la parole, pour un nombre de bandes fréquentielles variant de 2 à 16. Comme pour l'étude précédente (*« Sujet sain : parole masquée par du bruit »*, sect. 11.2, p. 235), les niveaux d'identification observés pour les voyelles et les consonnes sont plus élevés que pour les mots, cette dernière façon d'évaluer les performances perceptives étant en effet moins permissive. Cependant, contrairement à ce qui avait été observé avec une dégradation de la parole par du bruit blanc, les voyelles et les consonnes sont reconnues à des niveaux très similaires.

On observe l'apparition d'un maximum du gain audiovisuel dans la condition à 4 bandes fréquentielles ; au-delà de cette zone le gain audiovisuel est d'autant plus faible que le nombre de bandes fréquentielles est élevé (et donc que la parole est préservée). Comme dans le cas du masquage de la parole par un bruit blanc, il est intéressant de remarquer un phénomène d'inversion des gains audiovisuels associés aux mots entiers et aux consonnes/voyelles, cette inversion se produisant également pour la condition à 4 bandes fréquentielles : lorsque la résolution spectrale est meilleure (plus de bandes fréquentielles), les mots bénéficient d'un meilleur gain audiovisuel que les phonèmes pris individuellement, tandis que lorsqu'elle est moindre (moins de bandes fréquentielles), ce sont les voyelles et les consonnes individuelles qui bénéficient d'un gain audiovisuel plus élevé que les mots.

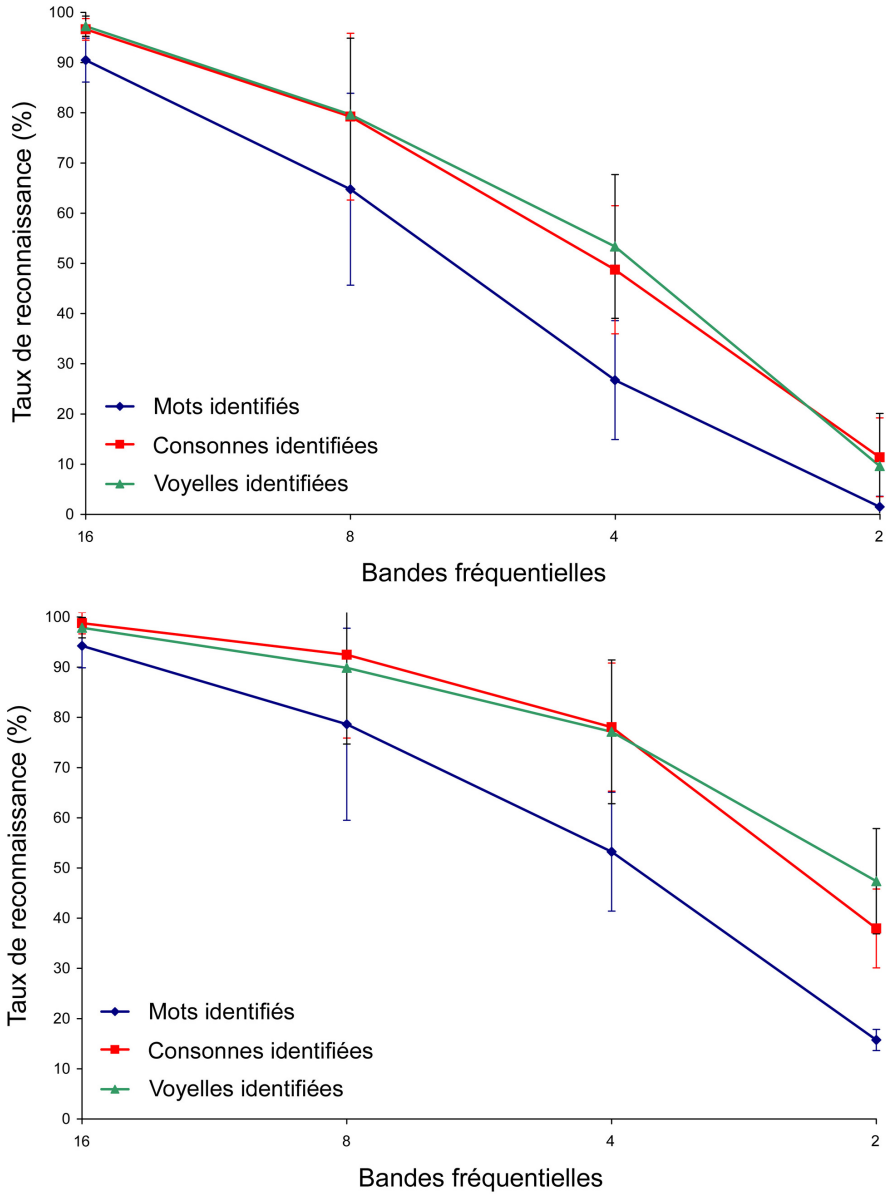


FIG. 11.8 – Performances de reconnaissance observées dans un groupe de 41 sujets normo-entendants exposés à de la parole auditive (en haut) et audiovisuelle (en bas) dégradée par une simulation acoustique du traitement effectué par l'implant cochléaire.

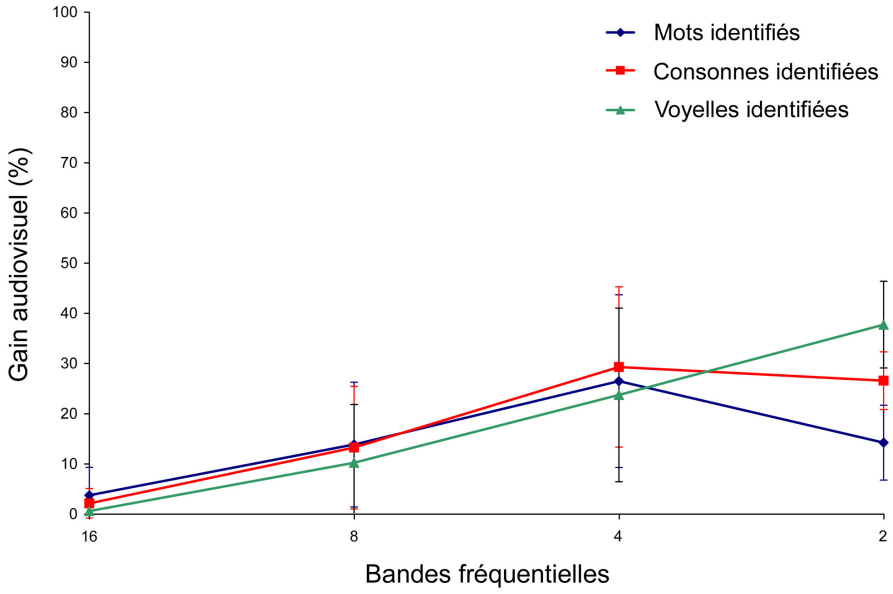


FIG. 11.9 – Gain audiovisuel observé dans un groupe de 41 sujets normo-entendants exposés à de la parole dégradée par une simulation acoustique du traitement effectué par l'implant cochléaire.

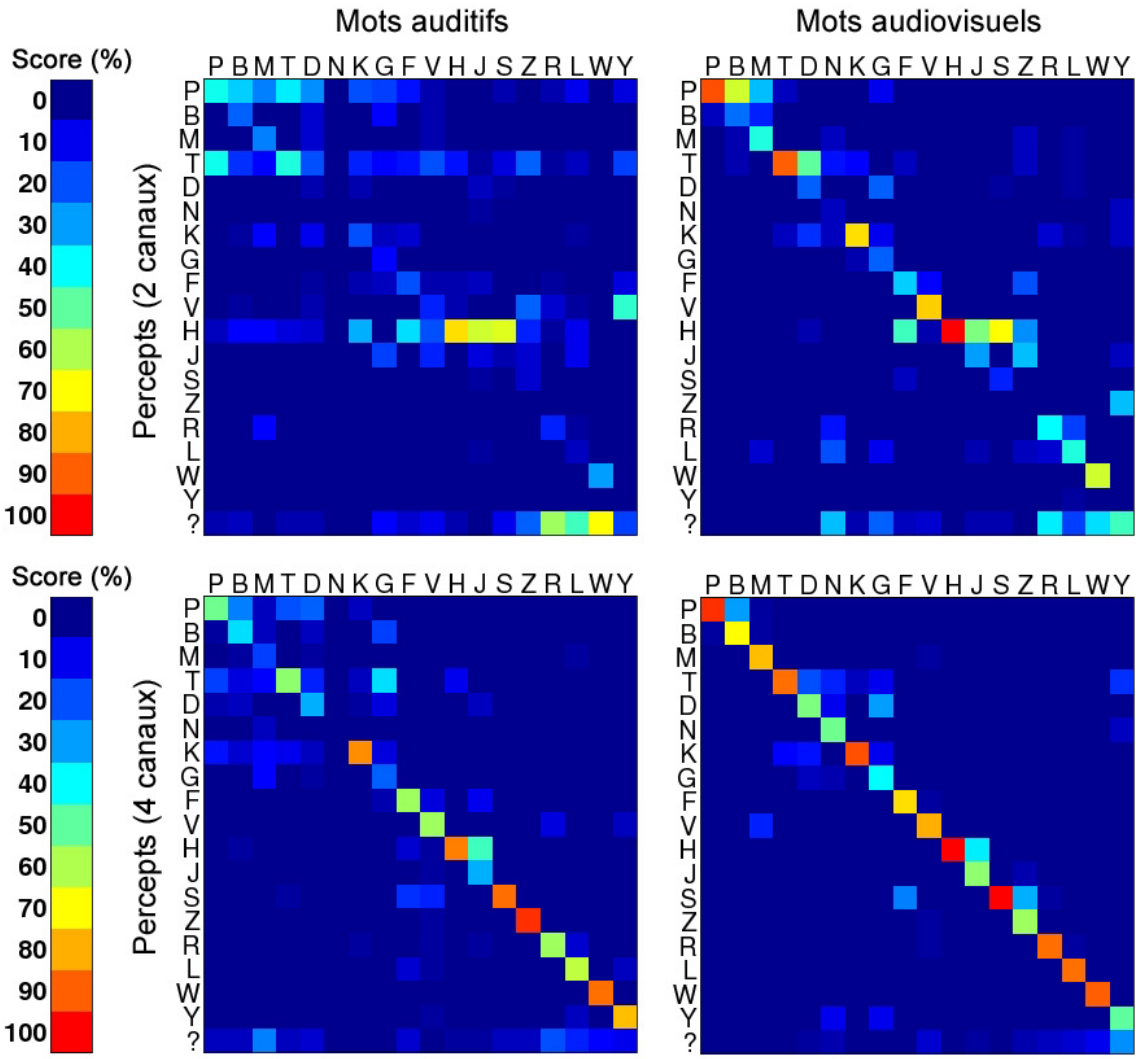


FIG. 11.10 – Analyse phonétique des confusions de consonnes réalisées par les sujets normaux sur des mots disyllabiques présentés en conditions auditive et audiovisuelle, avec une simulation du traitement sonore par l'implant cochléaire. Nombre d'électrodes simulées : 2 (n = 10) et 4 (n = 10).

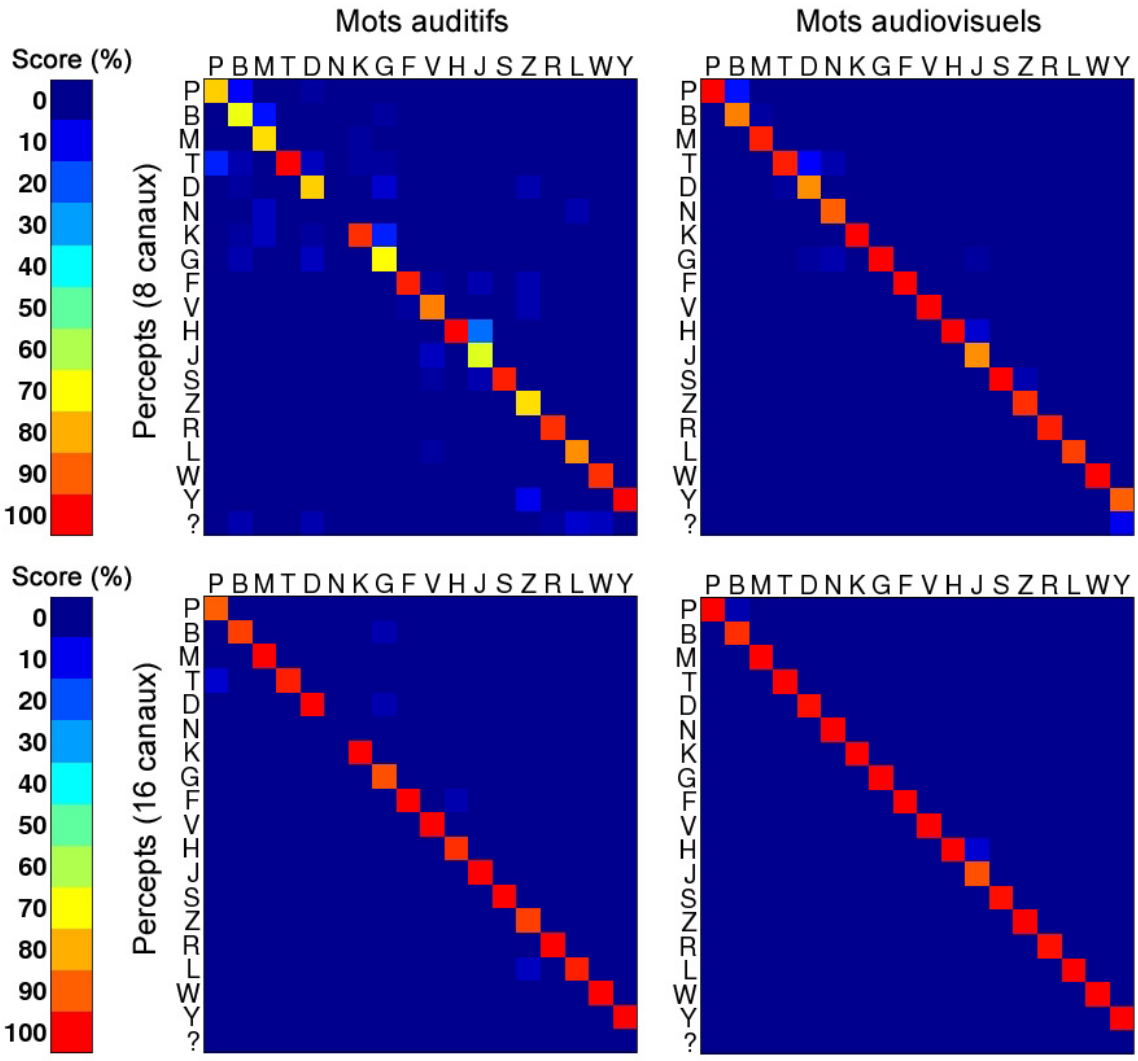


FIG. 11.11 – Analyse phonétique des confusions de consonnes réalisées par les sujets normaux sur des mots disyllabiques présentés en conditions auditive et audiovisuelle, avec une simulation du traitement sonore par l'implant cochléaire. Nombre d'électrodes simulées : 8 (n = 11) et 16 (n = 10).

11.3.3 Discussion

Cette étude indique des résultats comportementaux cohérents avec les études portant sur la perception audiovisuelle de la parole en condition dégradée, qui démontrent le bénéfice apporté par une présentation audiovisuelle (Sumby et Pollack, 1954; Miller et Nicely, 1955; Neely, 1956; Erber, 1969, 1975; Binnie *et al.*, 1974; Summerfield, 1979; MacLeod et Summerfield, 1987; Grant et Braida, 1991; Benoit *et al.*, 1994) (voir *« Améliorations perceptives en condition dégradée »*, § 3.2.2, p. 61).

Ici la présentation en condition audiovisuelle permet une amélioration perceptive allant jusqu'à plus de 20% suivant le nombre de bandes fréquentielles utilisées pour le filtrage. Nous retrouvons la présence d'un gain audiovisuel maximal pour un nombre de bandes fréquentielles

égales à 4, ce qui est à mettre en parallèle avec le maximum observé pour un rapport signal/bruit de -15 dB dans notre étude précédente (voir « *Sujet sain : parole masquée par du bruit* », sect. 11.2, p. 235), ainsi qu'avec la valeur de -12 dB indiquée dans la littérature (Ross *et al.*, 2007). Par ailleurs nous observons une diminution des performances auditives et audiovisuelles avec le nombre de bandes fréquentielles utilisées pour la simulation acoustique, ce qui est cohérent avec les observations rapportées pour des protocoles similaires dans la littérature (Shannon *et al.*, 1995; Loizou *et al.*, 1999).

De façon générale, on observe une amélioration audiovisuelle beaucoup plus restreinte lorsqu'une simulation acoustique d'implant cochléaire est utilisée que lorsque la parole est simplement masquée par un bruit blanc. Nous proposons une explication de cette différence dans le cadre de la distinction entre structure temporelle fine et enveloppe du signal (voir « *Rôle de l'enveloppe temporelle et de la structure fine* », § 9.2.2, p. 194). Tout d'abord, le masquage par du bruit blanc affecte de façon identique les différentes fréquences individuelles constitutives du spectre acoustique, sans altérer la corrélation temporelle naturelle qui existe entre ces différences composantes fréquentielles. Inversement, le traitement du signal opéré par l'implant cochléaire (équivalent au traitement acoustique utilisé dans notre simulation) détruit tous les indices associés à la structure temporelle fine du signal acoustique tout en préservant les informations d'enveloppe. Au sein d'une même bande spectrale, l'enveloppe temporelle des différentes fréquences individuelles est donc identique et équivalente à l'enveloppe temporelle globale du signal de bande (voir « *Simulation acoustique de l'implant cochléaire* », sect. 10.2.2, p. 218), au détriment des différences temporelles (structure fine) qui distinguent normalement les différentes fréquences de cette bande.

Par conséquent la simulation acoustique de l'implant cochléaire (ou le signal électrique délivré par un implant cochléaire), bien qu'elle préserve l'enveloppe temporelle globale dans de larges bandes spectrales, induit une détérioration partielle de l'enveloppe dans des bandes spectrales plus fines. Nous faisons l'hypothèse que cette détérioration temporelle auditive partielle a pour conséquence directe une détérioration de la cohérence audiovisuelle du signal de parole. Or l'intégration audiovisuelle est fortement dépendante de la cohérence audiovisuelle temporelle de la parole (voir « *Congruence temporelle de la parole* », § 3.3.1, p. 62). L'introduction d'une altération de la structure temporelle fine du signal acoustique, qui se traduit par une altération de la congruence temporelle audiovisuelle, est donc susceptible d'expliquer la baisse notable du gain audiovisuel observée dans le cas de la simulation acoustique d'implant cochléaire.

Ces résultats sont à mettre en parallèle avec les données expérimentales démontrant que la structure temporelle fine devient cruciale pour la compréhension de la parole dans un environnement bruité, aussi bien pour les sujets normo-entendants testés avec le vocoder (Nelson *et al.*, 2003; Qin et Oxenham, 2003; Stone et Moore, 2003; Zeng *et al.*, 2005; Fullgrabe *et al.*, 2006) que chez les sujets implantés cochléaires (Fu *et al.*, 1998; Friesen *et al.*, 2001) (voir également « *Rôle de l'enveloppe temporelle et de la structure fine* », § 9.2.2, p. 194). Ainsi nos observations tendent à indiquer que le rôle facilitateur de la structure temporelle fine interviendrait également lors de l'intégration audiovisuelle de la parole. Une discussion plus approfondie de l'apport de ces données sur la compréhension des mécanismes d'intégration audiovisuelle de la parole est fournie dans « *Intégration audiovisuelle chez le sujet implanté* », sect. 12.2, p. 262.

11.4 Entraînement perceptif avec une simulation d'implant

Cette étude a été motivée par plusieurs données expérimentales récentes démontrant que l'amélioration unimodale obtenue après un entraînement perceptif est meilleure pour un entraînement bimodal que pour un entraînement unimodal. Ainsi les données recueillies chez les sujets

normo-entendants montrent un surcroît d'amélioration des performances visuelles après un entraînement audiovisuel (Lehmann et Murray, 2005; Seitz *et al.*, 2006), tandis que les observations effectuées chez les patients implantés cochléaires montrent un surcroît d'amélioration visuelle (Lachs *et al.*, 2001) et auditive (Hodges *et al.*, 1999; Miyamoto *et al.*, 1999; Cullington *et al.*, 2000; Kirk *et al.*, 2000; Svirsky *et al.*, 2000) après un entraînement privilégiant la congruence audiovisuelle, comme le langage parlé complété (LPC) ou la parole audiovisuelle naturelle (voir également *« Effet de l'expérience sensorielle »*, § 9.3.2, p. 199).

Nous avons donc défini un protocole visant à reproduire chez des sujets normo-entendants l'entraînement perceptif subi par les patients sourds lors de l'utilisation quotidienne de leur implant cochléaire, afin de tester les effets de cet entraînement. Ainsi la présente étude a consisté à entraîner des sujets normo-entendants avec de la parole traitée par une simulation acoustique d'implant cochléaire. L'entraînement était réalisé de façon quotidienne, soit en condition auditive seule (sans l'image), soit en condition audiovisuelle. Des tests d'évaluation en condition auditive seule ont été pratiqués à intervalles réguliers avec des mots inconnus, afin de quantifier l'effet des différents types d'entraînement sur la perception auditive de la parole.

11.4.1 Matériel et méthodes

Sujets

Cette investigation a été effectuée dans le centre Cerveau et Cognition UMR CNRS 5549 à la faculté de médecine de Toulouse-Rangueil. Nous avons réalisé une étude prospective sur des sujets adultes ne présentant aucune déficience auditive. Le groupe de sujets est composé de 17 hommes entre 21 et 49 ans, avec une moyenne d'âge de 27.8 ans ($\pm$ 6.6 ans). Tous les sujets étaient entièrement naïfs vis-à-vis du protocole expérimental, aucun d'entre eux n'ayant jamais été exposé auparavant aux stimuli utilisés.

Stimuli

Sélection des stimuli : Les stimuli sont tous des mots dissyllabiques faisant partie des listes normalisées de Fournier. Ces mots ont été répartis en deux listes de 40 mots chacune (L1 et L2) et une liste de 30 mots (L3). Ces listes ont été équilibrées en fréquence d'utilisation lexicale, comptabilisée en nombre d'occurences par million de mots d'après la base de données Brulex (Content *et al.*, 1990), en structure phonétique (répartition voyelles antérieures/voyelles postérieures) et en structure syllabique (CV/CVC/CCV). L'enregistrement des prononciations de chaque mot s'est déroulé suivant les modalités décrites dans *« Génération des stimuli »*, § 10.2.1, p. 218.

Génération des stimuli : Les sons ont été travaillés sous Matlab 6.5 sur un processeur Pentium 4 2.4 GHz suivant une procédure SSSN (speech spectrum shaped noise). Ils ont été analysés dans différentes bandes spectrales à l'aide de filtres elliptiques IIR du sixième ordre suivant la carte tonotopique de Greenwood (Greenwood, 1990) ; seules les fréquences comprises entre 250 et 8000 Hz ont été retenues, avec un découpage en quatre bandes suivant les fréquences de coupure 250, 709, 1676, 3713 et 8000 Hz.

Les opérations de traitement de signal utilisées pour la réalisation de ces stimuli sont identiques à celles utilisées dans l'étude précédente, et sont décrites dans *« Simulation acoustique de l'implant cochléaire »*, sect. 10.2.2, p. 218. On pourra consulter *« Influence de la profondeur d'implantation »*, § 9.2.3, p. 196 pour plus de détails sur la procédure SSSN et pour des élements de discussion sur la validité d'utilisation de la carte tonotopique de Greenwood dans le cadre d'une simulation acoustique d'implant cochléaire.

Les stimuli ont été montés ensemble sous forme de films vidéo au format MPEG2 avec Adobe Premiere Pro 7.0 sur un processeur Pentium 4 2.4 GHz. Les listes L1 et L2, utilisées en version auditive seule ou audiovisuelle en fonction des sujets, ont été réservées à l'entraînement perceptif et mélangées dans trois ordres aléatoires différents, fournissant ainsi six listes d'entraînement. La liste L3, utilisée uniquement en version auditive seule, a été réservée à l'évaluation perceptive et mélangée dans trois ordres aléatoires différents, fournissant ainsi trois listes d'évaluation.

Ordre de présentation : Tous les sujets ont subi un entraînement perceptif pendant trois semaines consécutives. Chaque semaine, pendant cinq jours consécutifs (du lundi au vendredi), les sujets ont passé le test d'entraînement en début de journée. Une partie des sujets (n = 7) a été entraînée en condition auditive seule et une autre partie (n = 7) en condition audiovisuelle. A la fin de chaque semaine (le vendredi en fin de journée), tous les sujets ont passé le test d'évaluation en condition auditive seule, quelle que soit la condition d'entraînement (auditive ou audiovisuelle). En outre, tous les sujets ont passé le test d'évaluation au moins une semaine avant le début de l'entraînement perceptif, afin de contrôler leurs performances dans un état naïf, puis deux semaines après la fin de l'entraînement perceptif, afin de tester l'effet de l'entraînement perceptif à moyen terme.

Pour finir, un sous-groupe de sujets (n = 3) a été testé uniquement avec les listes d'évaluation (une fois par semaine), sans aucun entraînement perceptif, afin de quantifier l'effet d'entraînement intrinsèque aux listes d'évaluation.

Procédure

Les stimuli ont été présentés sur un ordinateur portable avec un moniteur de dimensions 20.7×33.1 cm. L'image était présentée au centre de l'écran avec une dimension fixe de 20.7×27.6 cm, le reste de l'écran demeurant noir. Les sujets étaient assis confortablement en face de l'écran, à une distance de lecture normale d'environ 65 cm, ce qui correspond à un diamètre angulaire d'environ 30°. Le son était délivré à travers un casque d'écoute avec une faible déformation spectrale, de telle sorte que le niveau sonore moyen atteigne 65 dB.

Les participants ont tous été testés individuellement. Ils ont reçu pour seules instructions de regarder l'image et d'écouter le son pendant toute la durée de présentation de chaque stimulus, puis de répéter à voix haute ce qu'ils avaient perçu. L'expérimentateur était assis dans la pièce pendant toute la durée de l'expérience pour confirmer que les sujets regardaient bien l'écran et pour consigner par écrit leurs réponses verbales.

11.4.2 Résultats

Listes d'entraînement

Les résultats montrent une très nette amélioration comportementale dès la fin de la première semaine d'entraînement (voir fig. 11.12, p. 252), aussi bien en condition auditive (gain de 12.5% sur la semaine) qu'en condition audiovisuelle (gain de 20% sur la semaine), les améliorations étant statistiquement significatives dans les deux cas (T5-T1 ; $p < 0.05$). Cette amélioration se poursuit pendant la deuxième semaine pour l'entraînement auditif (T10-T6 ; $p < 0.05$), tandis qu'elle se stabilise pour l'entraînement audiovisuel (T10-T6 ; $p > 0.05$). A partir de la troisième semaine l'ensemble des sujets voient leurs performances se stabiliser, quelle que soit la condition d'entraînement (T15-T11 ; $p > 0.05$).

On peut remarquer que les marges de progression globales (T15-T1) sont tout à fait similaires, avec une amélioration de 28.6% lors de l'entraînement audiovisuel, à comparer avec une amélioration de 29.6% lors de l'entraînement auditif (pas de différence ; $p > 0.05$). Conformément à ce que l'on pouvait attendre, les performances des sujets testés en condition audiovisuelle sont systématiquement meilleures que celles des sujets testés en condition auditive seule, une différence de 30% séparant les scores audiovisuels des scores auditifs.

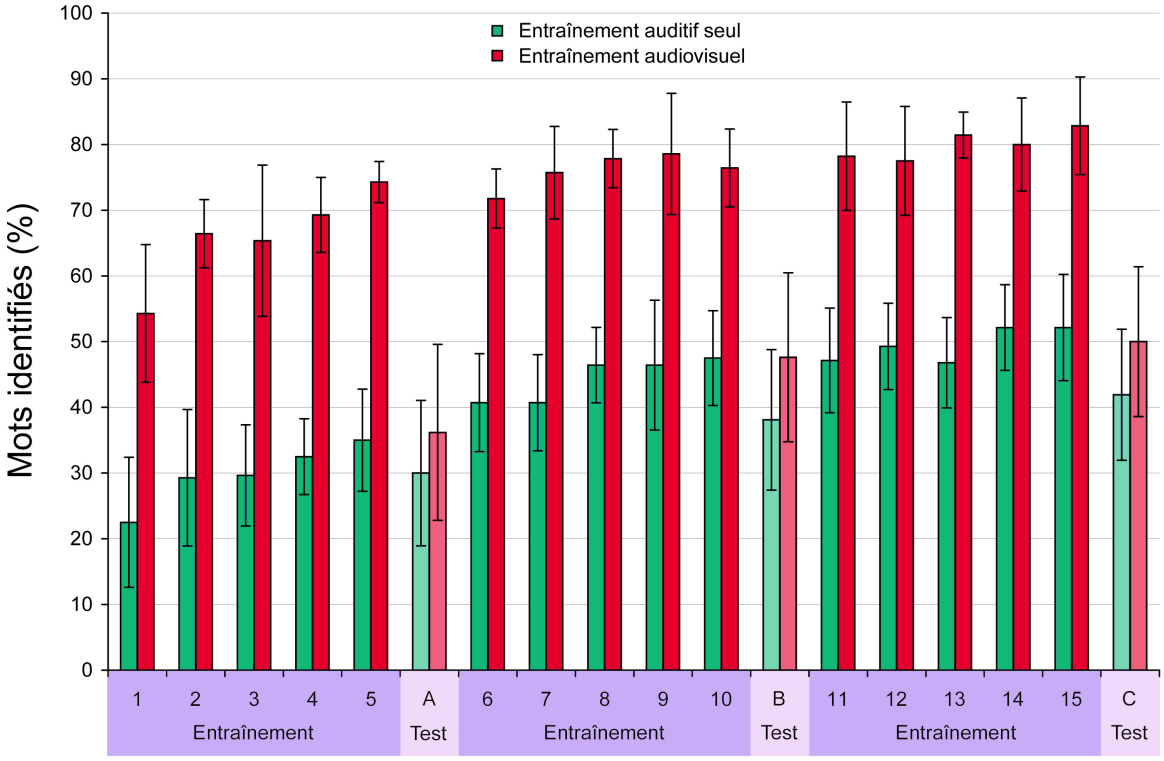


FIG. 11.12 – Compréhension de la parole traitée par une simulation d'implant cochléaire lors d'un entraînement perceptif en conditions auditive seule (n = 7) et audiovisuelle (n = 7) chez des sujets normo-entendants. Les tests ont tous été réalisés en condition auditive seule.

Listes d'évaluation

Lors de l'évaluation auditive, le groupe de sujets n'ayant pas subi d'entraînement perceptif présente une légère progression de 13% (comparaison entre pré-test et test C) sur l'ensemble des trois semaines (voir fig. 11.13, p. 253), démontrant ainsi un léger effet d'entraînement intrinsèque à la liste d'évaluation. Toutefois cette progression intrinsèque reste très largement inférieure aux progressions constatées dans les groupes de sujets entraînés (voir fig. 11.13, p. 253), pour lesquelles on observe une progression de 32% en auditif et de 40% en audiovisuel (différence significative avec la condition sans entraînement ; $p < 0.05$).

Il est particulièrement intéressant de remarquer que lors du test d'évaluation en condition auditive seule, les sujets entraînés en condition audiovisuelle obtiennent de meilleures performances globales en comparaison avec les sujets entraînés en condition auditive seule (voir fig. 11.13, p. 253), la différence s'étalant entre 5 et 8% en fonction du moment du test. Cette différence tend à s'accentuer au cours de l'entraînement et persiste même deux semaines après

la fin de l'entraînement perceptif.

Cependant les différences de performances observées lors des évaluations en fonction de la modalité d'entraînement n'apparaissent pas comme statistiquement significatives. Comme ces différences se répètent de façon très consistante quelle que soit la date de l'évaluation, nous en déduisons que le groupe expérimental utilisé n'est pas assez important pour faire ressortir l'effet recherché. Un approfondissement de ce paradigme expérimental est donc nécessaire pour dresser des conclusions pertinentes.

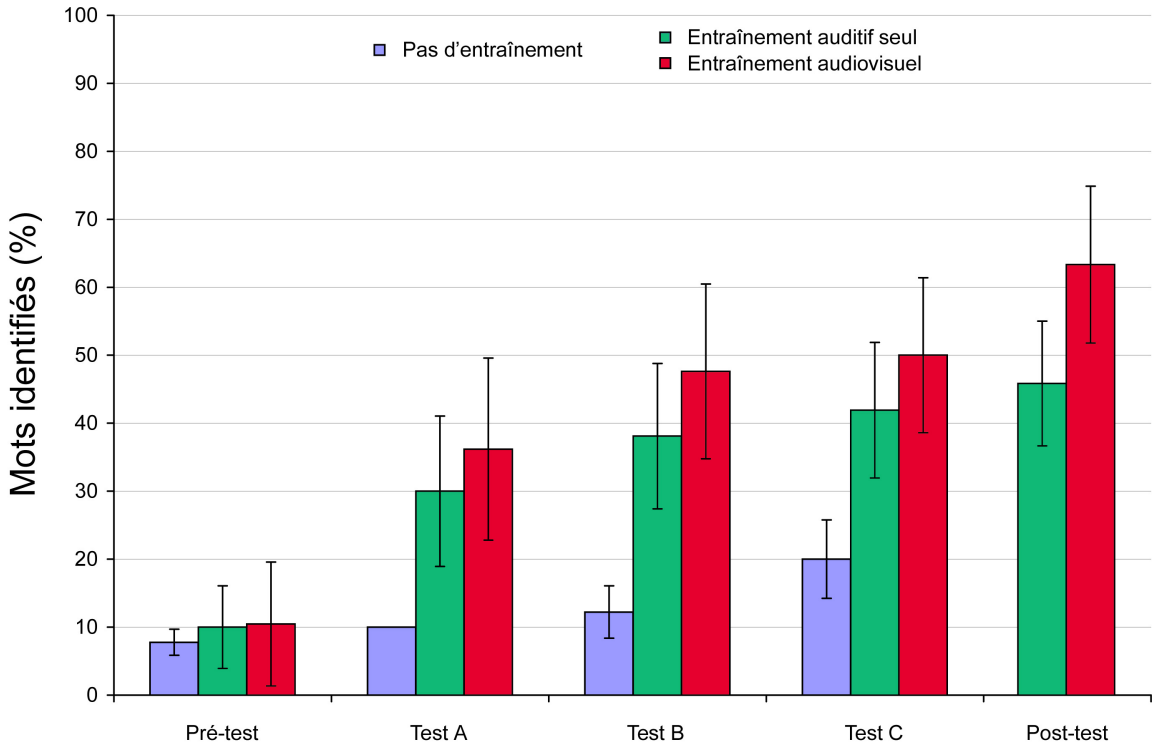


FIG. 11.13 – Compréhension de la parole traitée par une simulation d'implant cochléaire lors des tests auditifs réalisés pour les différentes conditions d'entraînement des sujets normo-entendants.

11.4.3 Discussion

Nos résultats montrent le rôle primordial joué par l'intégration audiovisuelle lors de la perception de la parole, en particulier dans la résolution des ambigüités et dans l'apprentissage perceptif dans des conditions de dégradation des informations auditives, comme c'est le cas chez les porteurs d'un implant cochléaire. Nous démontrons ici la répercussion directe d'un apprentissage en condition bimodale sur les performances en condition unimodale, ce qui confirme l'influence bénéfique d'un entraînement sensoriel audiovisuel dans la perception de la parole en condition auditive seule.

Nous rappelons sur ce point les données recueillies chez les sujets normo-entendants et les patients implantés cochléaires, qui montrent un surcroît d'amélioration des performances unimodales après un entraînement bimodal (Hodges *et al.*, 1999; Miyamoto *et al.*, 1999; Cullington *et al.*, 2000; Kirk *et al.*, 2000; Svirsky *et al.*, 2000; Lachs *et al.*, 2001; Lehmann et Murray, 2005; Seitz *et al.*, 2006). Par ailleurs une étude très récente (Rosenblum *et al.*, 2007), réalisée

avec des sujets normo-entendants, a même démontré qu'on pouvait observer une amélioration de la compréhension auditive de la parole dégradée à la suite d'un entraînement perceptif réalisé uniquement en condition visuelle (lecture labiale), l'effet étant moins marqué si des locuteurs différents sont utilisés pour les stimulations auditive et visuelle. De tels résultats confirment de façon spectaculaire le rôle primordial joué par la composante visuelle de la parole dans l'apprentissage et la réhabilitation auditive de la parole.

Notre étude trouve une application clinique dans le cas de la réhabilitation orthophonique des patients ayant subi une implantation cochléaire. En effet, l'implant cochléaire ne restituant qu'imparfaitement la richesse spectrale de l'environnement sonore des patients, ces derniers se trouvent ainsi placés dans une situation où les informations auditives sont dégradées, ce qui devrait leur permettre de tirer un bénéfice particulier de l'amélioration perceptive rendue possible par l'intégration audiovisuelle de la parole. Nous fournissons ici des éléments de preuve expérimentale indiquant qu'une réhabilitation orthophonique mettant particulièrement l'accent sur une stimulation audiovisuelle est fortement susceptible d'induire chez les patients implantés une meilleure et plus rapide progression des performances auditives qu'une réhabilitation basée sur une stimulation uniquement auditive.

Bibliographie

Aloufy, S., Lapidot, M. et Myslobodskym, M. (1996). Differences in susceptibility to the "blending illusion" among Native Hebrew and English speakers. *Brain Lang.*, 53(1):51–57.

Benoit, C., Mohamadi, T. et Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *J. Speech Hear. Res.*, 37(5):1195–1203.

Binnie, C. A., Montgomery, A. A. et Jackson, P. L. (1974). Auditory and visual contributions to the perception of consonants. *J. Speech Hear. Res.*, 17(4):619–630.

Content, A., Mousty, P. et Radeau, M. (1990). Brulex. une base de données lexicales informatisée pour le français écrit et parlé. *Ann. Psychol.*, 90:551–566.

Cullington, H., Hodges, A. V., Butts, S. L., Dolan-Ash, S. et Balkany, T. J. (2000). Comparison of language ability in children with cochlear implants placed in oral and total communication educational settings. *Ann. Otol. Rhinol. Laryngol. Suppl.*, 185:121–123.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *J. Speech Hear. Res.*, 12(2):423–425.

Erber, N. P. (1975). Auditory-visual perception of speech. *J. Speech Hear. Disord.*, 40(4):481–492.

Friesen, L. M., Shannon, R. V., Baskent, D. et Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels : comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.*, 110(2):1150–1163.

Fu, Q. J., Shannon, R. V. et Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition : acoustic and electric hearing. *J. Acoust. Soc. Am.*, 104(6):3586–3596.

Fullgrabe, C., Berthommier, F. et Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.*, 211(1-2):74–84.

Goh, W. D., Pisoni, D. B., Kirk, K. I. et Remez, R. E. (2001). Audio-visual perception of sinewave speech in an adult cochlear implant user : a case study. *Ear Hear.*, 22(5):412–419.

Grant, K. W. et Braida, L. D. (1991). Evaluating the articulation index for audiovisual input. *J. Acoust. Soc. Am.*, 89:2952–2960.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species - 29 years later. *J. Acoust. Soc. Am.*, 87:2592–2605.

Hodges, A. V., Dolan Ash, M., Balkany, T. J., Schloffman, J. J. et Butts, S. L. (1999). Speech perception results in children with cochlear implants : contributing factors. *Otolaryngol. Head Neck Surg.*, 121(1):31–34.

Irwin, J. R., Whalen, D. H. et Fowler, C. A. (2006). A sex difference in visual influence on heard speech. *Percept. Psychophys.*, 68(4):582–592.

Johnson, F. M. et Leslie, H. H. (1988). Sex differences in lipreading. *Bull. Psychon. Soc.*, 26(2):106–108.

Kirk, K. I., Pisoni, D. B. et Miyamoto, R. T. (2000). Lexical discrimination by children with cochlear implants : Effects of age at implantation and communication mode. *In* Waltzman, S. et Cohen, N., éditeurs : *Proceedings of the Vth International Cochlear Implant Conference*. New York : Thieme Medical Publishers.

Lachs, L., Pisoni, D. B. et Kirk, K. I. (2001). Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants : a first report. *Ear Hear.*, 22(3):236–251.

Lehmann, S. et Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.*, 24(2):326–334.

Loizou, P. C., Dorman, M. et Tu, Z. (1999). On the number of channels needed to understand speech. *J. Acoust. Soc. Am.*, 106(4 Pt 1):2097–2103.

MacLeod, A. et Summerfield, A. Q. (1987). Quantifying the contribution of vision to speech perception in noise. *Br. J. Audiol.*, 21(2):131–141.

Miller, G. A. et Nicely, P. (1955). An analysis of perceptual confusions among some english consonants. *J. Acoust. Soc. Am.*, 27(2):338–352.

Miyamoto, R. T., Kirk, K. I., Svirsky, M. A. et Sehgal, S. T. (1999). Communication skills in pediatric cochlear implant recipients. *Acta Otolaryngol.*, 119(2):219–224.

Neely, K. K. (1956). Effect of visual factors on the intelligibility of speech. *J. Acoust. Soc. Am.*, 28:1275–1277.

Nelson, P. B., Jin, S. H., Carney, A. E. et Nelson, D. A. (2003). Understanding speech in modulated interference : cochlear implant users and normal-hearing listeners. *J. Acoust. Soc. Am.*, 113(2):961–968.

Qin, M. K. et Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.*, 114(1):446–454.

Risberg, A. et Lubker, J. (1978). Prosody and speechreading. *Speech Trans. Lab. Q. Prog. Stat. Rep.*, 4:1.

Rosenblum, L. D., Miller, R. M. et Sanchez, K. (2007). Lip-read me now, hear me better later : cross-modal transfer of talker-familiarity effects. *Psychol. Sci.*, 18(5):392–396.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. et Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb. Cortex*, 17(5):1147–1153.

Seitz, A. R., Kim, R. et Shams, L. (2006). Sound facilitates visual learning. *Curr. Biol.*, 16(14):1422–1427.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonsky, J. et Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270:303–304.

Stone, M. A. et Moore, B. C. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *J. Acoust. Soc. Am.*, 114(2):1023–1034.

Sumby, W. H. et Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.*, 26:212–215.

Summerfield, A. Q. (1979). Use of visual information for phonetic perception. *Phonetica*, 36:314–331.

Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B. et Miyamoto, R. T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychol. Sci.*, 11(2):153–158.

Watson, C. S., Qiu, W. W., Chamberlain, M. M. et Li, X. (1996). Auditory and visual speech perception : confirmation of a modality-independent source of individual differences in speech recognition. *J. Acoust. Soc. Am.*, 100(2 Pt 1):1153–1162.

Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C. et Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. U.S.A.*, 102(7):2293–2298.

Chapitre 12

Etudes comportementales chez le patient implanté

Sommaire

12.1 Perception de la parole chez le sujet implanté 258
12.1.1 Matériel et méthodes 258
12.1.2 Résultats 259
12.1.3 Discussion 261
12.2 Intégration audiovisuelle chez le sujet implanté 262
12.2.1 Matériel et méthodes 262
12.2.2 Résultats et discussion 264
12.2.3 Conclusion 266
12.3 Lecture labiale chez le sujet implanté 268
12.3.1 Matériel et méthodes 268
12.3.2 Résultats et discussion 269
12.4 Effet McGurk chez le sujet implanté 271
12.4.1 Matériel et méthodes 271
12.4.2 Résultats 273
12.4.3 Discussion 278
Bibliographie 285

Introduction

Afin de mieux comprendre l'évolution des mécanismes sensoriels impliqués dans la perception audiovisuelle de la parole après une implantation cochléaire, nous avons défini plusieurs études comportementales destinées à apporter des éléments de réponse à cette problématique spécifique.

Notre première étude (voir *« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258 et *« Intégration audiovisuelle chez le sujet implanté »*, sect. 12.2, p. 262) porte sur le suivi longitudinal postimplantation à long terme de la compréhension de la parole en conditions visuelle, auditive et audiovisuelle (n_{IC}= 97). Des comparaisons

sont effectuées avec des sujets normo-entendants testés dans les conditions les plus similaires possibles, en l'occurence en lecture labiale (n_{NH}= 42) puis en stimulation auditive et audiovisuelle avec une dégradation du son par du bruit blanc (n_{NH}= 80) et par une simulation acoustique d'implant cochléaire (n_{NH}= 41).

Cette étude longitudinale comparative est destinée à explorer en première approche les contributions respectives de l'audition et de la vision lors de l'intégration audiovisuelle de la parole, celle-ci permettant une nette amélioration perceptive (Kaiser *et al.*, 2003; Geers *et al.*, 2003; Bergeson *et al.*, 2003, 2005; Moody-Antonio *et al.*, 2005). Par ailleurs, les implantés cochléaires présentant des performances supranormales en lecture labiale (Giraud *et al.*, 2001; Bergeson *et al.*, 2003, 2005), nous approfondissons la question au travers d'une analyse longitudinale détaillée de la lecture labiale chez les patients implantés et les sujets normo-entendants, en effectuant une distinction entre hommes et femmes (voir *« Lecture labiale chez le sujet implanté »*, sect. 12.3, p. 268).

Enfin, dans une dernière étude, nous utilisons des stimuli de parole audiovisuelle incongruente de type McGurk (voir *« Effet McGurk chez le sujet implanté »*, sect. 12.4, p. 271) afin d'effectuer une analyse fine des processus d'intégration phonétique des informations auditives et visuelles chez les patients implantés (n_{NH}= 33) et les sujets normo-entendants (n_{IC}= 39).

12.1 Perception de la parole chez le sujet implanté

12.1.1 Matériel et méthodes

Sujets

Nous avons réalisé une analyse rétrospective portant sur les données comportementales orthophoniques de 97 patients adultes présentant une surdité profonde bilatérale postlinguale (perte auditive supérieure à 90 dB) et ayant subi une implantation cochléaire unilatérale. Le groupe de sujets est composé de 60 femmes et 37 hommes entre 19 et 82 ans, avec une moyenne d'âge de 56.2 ans (± 17.0 ans). Les étiologies sont très variables dans le groupe de sujets, principalement infectieuses et dégénératives (méningite, otite chronique, otosclérose, neurinome acoustique, otospongiose, etc.) La plupart des surdités sont progressives, ce qui rend difficile une évaluation fiable de la durée de surdité, approximativement étalée entre 1 et 57 ans sur le groupe de sujets, avec une moyenne de 22 ans. Tous les patients ont reçu un implant cochléaire Nucleus (CI-22 ou CI-24), avec l'une des stratégies de codage MPEAK, SPEAK ou ACE. Les critères de sélection pour l'implantation cochléaire requièrent des résultats d'audiométrie vocale en champ libre inférieurs à 30% pour des mots et des phrases présentés dans les meilleures conditions d'aide auditive possible avec des prothèses auditives conventionnelles.

Stimuli

Les stimuli utilisés sont des mots disyllabiques provenant des listes normalisées de mots de Fournier, regroupés par séries de 10 mots.

Procédure

Tous les patients ont été testés par la même orthophoniste dans le service ORL du CHU Toulouse-Purpan. Le test est réalisé tous les mois dans le premier trimestre suivant la mise en marche de l'implant, puis tous les trois mois dans la première année, puis enfin tous les ans

après la première année. Pour chaque séance, trois séries de dix mots disyllabiques sont testées en lecture labiale seule (V), avec l'implant cochléaire seul (A), puis en condition audiovisuelle complète (AV). Les mots sont prononcés par l'orthophoniste face au patient. La stimulation auditive, lorsqu'elle est présente, consiste à réaliser une audiométrie vocale en champ libre dans le silence à un niveau sonore de 65 dB.

12.1.2 Résultats

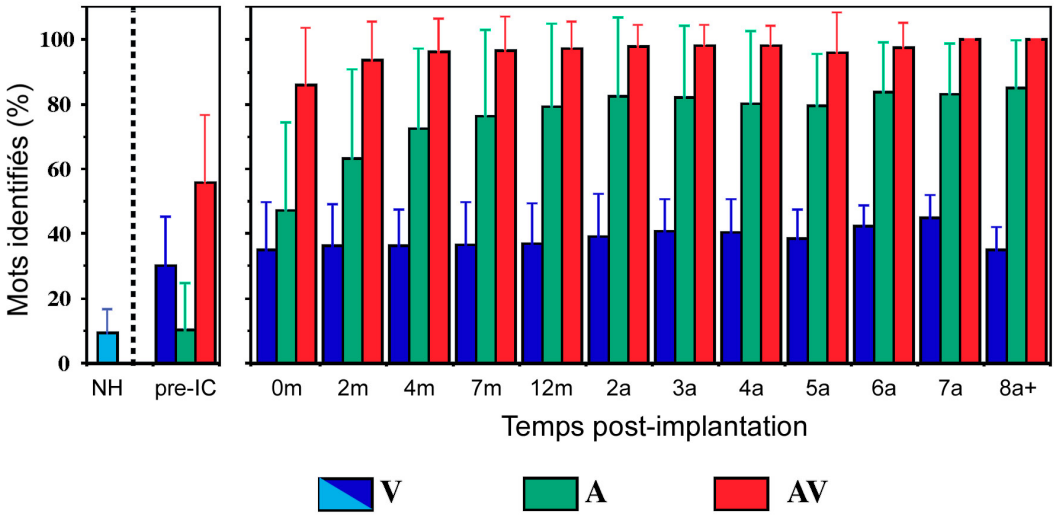


FIG. 12.1 – Niveaux de reconnaissance de la parole en conditions visuelle, auditive et audiovisuelle lors du suivi longitudinal d'un groupe de 97 patients implantés cochléaires. Pour référence, la partie gauche du graphique indique les scores perceptifs des mêmes patients avant leur implantation cochléaire, testés avec des prothèses acoustiques conventionnelles (pre-IC), ainsi que les scores de lecture labiale observés dans un groupe de 42 sujets contrôles (NH ; cf. *« Lecture labiale chez le sujet sain »*, sect. 11.1, p. 232).

Performances avant implantation

Les personnes sourdes sélectionnées pour l'implantation cochléaire sont testées en audiométrie vocale dans un délai inférieur à un an avant l'intervention chirurgicale, le délai moyen étant d'environ 6 mois avant implantation. Les résultats de l'audiométrie vocale pré-implantation (voir fig. 12.1, p. 259) montrent une compréhension très difficile en audition seule (10,4 ± 14,2%). Le test en modalité audiovisuelle montre une nette amélioration de la compréhension en comparaison avec l'audition seule (55.8 ± 21.0%, $p < 0.05$) ; la compréhension reste toutefois malaisée. Enfin la compréhension en lecture labiale seule est assez bonne et dans tous les cas nettement supérieure aux performances des sujets normo-entendants testés avec le même locuteur (30.1 ± 15.1% > 9.4 ± 7.1%, $p < 0.05$).

Evolution de l'audition

Le groupe de patients implantés montre une amélioration significative de la compréhension en audition seule (voir fig. 12.1, p. 259) entre la période pré-implantation et la mise en marche de l'implant cochléaire (47.1 ± 27.3%, $p < 0.05$). La compréhension auditive de la parole s'améliore

de façon significative depuis l'activation de l'implant jusqu'au septième mois suivant l'implantation ($p < 0.05$), avant d'atteindre un plateau (performance moyenne de 81%) ne montrant plus de progression significative au niveau de la population complète (bien qu'à un niveau individuel, pour de nombreux patients, la progression des performances auditives continue au-delà du septième mois post-implantation).

Evolution de la compréhension audiovisuelle

Lors de la mise en marche de l'implant cochléaire, les sujets implantés montrent une nette augmentation ($p < 0.05$) de la compréhension pour une stimulation audiovisuelle ($86.0 \pm 17.4\%$), en comparaison avec les stimulations en lecture labiale ($35.1 \pm 14.7\%$) ou en audition seule ($47.1 \pm 27.3\%$) (voir fig. 12.1, p. 259). Comme dans le cas de l'audition seule, on observe une progression significative de la compréhension audiovisuelle de la parole avec le temps, ce jusqu'au deuxième mois suivant l'implantation ($p < 0.05$), où les performances audiovisuelles sont excellentes ($94.0 \pm 12.0\%$).

De la même façon que pour l'audition seule, les performances audiovisuelles lors de l'activation de l'implant sont supérieures à celles observées avant implantation ($86.0 \pm 17.4\% > 55.8 \pm 21.0\%$, $p < 0.05$). Comme seules les informations sensorielles auditives diffèrent entre ces deux périodes, on peut faire l'hypothèse que l'augmentation brusque des performances audiovisuelles est uniquement liée à l'amélioration de la compréhension auditive. Cette hypothèse tend a être confirmée par une analyse de sous-groupe portant sur les patients ne présentant aucune amélioration auditive à la mise en marche de l'implant ($n = 14$) : on observe en effet que ces patients ne présentent aucune augmentation significative du gain audiovisuel (0.54 vs. 0.62, différence non significative avec $p = 0.23$).

Evolution de la lecture labiale

Les patients testés à T0, soit moins de deux semaines après la mise en marche de l'implant, présentent une compréhension en lecture labiale (voir fig. 12.1, p. 259) toujours élevée ($35.1 \pm 14.7\%$) et comparable au niveau de compréhension mesuré avant l'implantation ($30.1 \pm 15.1\%$, différence non significative avec $p = 0.62$). Ces performances en lecture labiale restent stables quelle que soit la période post-implantation considérée (jusqu'à au moins 8 ans après la mise en marche de l'implant ; $p > 0.05$).

On n'observe aucune corrélation entre la compréhension en lecture labiale seule et la compréhension en audition seule ($r^2 = 0.001$, $p = 0.77$).

Les performances visuelles obtenues par les patients à la mise en marche de l'implant ne montrent aucune corrélation avec la durée de surdité ($r^2 = 0.001$, $p = 0.76$). Ce résultat doit toutefois être nuancé en raison du manque de fiabilité de l'évaluation de la durée des surdités, celles-ci étant la plupart du temps progressives avec des échelles de temps pouvant atteindre plusieurs dizaines d'années. Afin d'obtenir plus d'informations sur un éventuel lien entre durée de surdité et compréhension en lecture labiale, nous avons isolé un sous-groupe de patients ayant subi une surdité brusque (méningite ou traumatisme crânien) dans un délai inférieur à un an avant l'implantation. Malgré un faible effectif ($n = 3$), les scores en lecture labiale observés dans ce sous-groupe à la mise en marche de l'implant semblent similaires à ceux observés dans la population globale des 97 patients (20, 30 et 45%).

Pour préciser cette observation nous avons pris en compte tous les patients qui avaient subi une implantation cochléaire après notre étude prospective (ils n'avaient donc pas été inclus dans les analyses de groupe), et qui avaient également subi une surdité brusque dans un délai inférieur à un an avant implantation. 5 patients supplémentaires correspondaient à ce critère, ce qui nous

a permis d'effectuer une comparaison sur un total de 8 personnes. Dans ce groupe restreint, la compréhension en lecture labiale au moment de la mise en marche de l'implant se maintient à un niveau nettement supérieur à celui observé chez les sujets normo-entendants ($27.5 \pm 10.7\% > 9.4 \pm 7.1\%$, $p < 0.05$) tout en restant comparable à la population globale des 97 patients implantés ($p > 0.05$; différence non significative).

12.1.3 Discussion

Perception auditive de la parole chez les patients implantés

Conformément aux résultats rapportés dans la littérature scientifique, nos données cliniques indiquent une rapide récupération des performances auditives après l'implantation cochléaire qui permet d'atteindre un niveau de reconnaissance auditive de 8 mots sur 10 sans contexte grammatical ou sémantique. Ces observations s'accordent avec celles déjà rapportées dans le cas des enfants sourds prélinguaux et des adultes sourds postlinguaux, indiquant une rapide progression auditive après implantation (Pelizzone *et al.*, 1999; Hamzavi *et al.*, 2003; Bergeson *et al.*, 2003) (voir *« Perception auditive de la parole après implantation »*, § 9.1.1, p. 192).

Ces données fournissent également des informations importantes sur la fenêtre temporelle dans laquelle se produit cette récupération perceptive, en montrant une stabilisation du niveau de compréhension de la parole auditive dans la première année suivant la mise en marche de l'implant cochléaire (à partir du septième mois). Ceci suggère fortement l'intervention de mécanismes de coopération audiovisuelle (voir ci-après) qui contribueraient à cette progression rapide et efficace.

Perception visuelle de la parole chez les patients implantés

Nos données cliniques indiquent : 1° que l'acquisition de compétences perceptives en lecture labiale est susceptible de se faire rapidement après une déprivation sensorielle auditive, dans une échelle de temps de quelques mois ; 2° qu'après la mise en marche de l'implant, cette spécialisation perceptive reste stable sur le long terme, sans être affectée par une fréquente et durable exposition aux signaux auditifs artificiels délivrés par l'implant cochléaire.

Ces données ne semblent toutefois pas indiquer une progression des performances visuelles à la suite de l'implantation, comme cela avait pu être observé dans d'autres études similaires (Tyler *et al.*, 1997; Giraud *et al.*, 2001; Bergeson *et al.*, 2003, 2005) (voir *« Perception visuelle de la parole après implantation »*, § 9.1.2, p. 193). Cette non-significativité pourrait être liée à divers facteurs comme l'hétérogénéité de la population des patients implantés, tout particulièrement concernant l'étiologie de la surdité. On pourra se référer à *« Lecture labiale chez le sujet implanté »*, sect. 12.3, p. 268, pour une discussion détaillée des facteurs longitudinaux influant sur les performances de lecture labiale des patients implantés.

La conservation d'un niveau supra-normal de lecture labiale par les patients implantés pourrait être liée à des stratégies audiovisuelles compensatrices intervenant en particulier dans des situations de dégradation supplémentaire des informations sonores, comme par exemple dans un environnement bruité. En effet les performances auditives des patients implantés sont très facilement affectées par un bruit ambiant, ce qui nécessite la mise en place de stratégies perceptives compensatrices (voir discussion ci-après).

Perception audiovisuelle de la parole chez les patients implantés

De la même façon que pour la compréhension auditive de la parole, les données cliniques montrent une rapide récupération des performances de reconnaissance audiovisuelle après l'implantation cochléaire, les patients atteignant en moyenne un niveau de compréhension de 19 mots

sur 20 sans contexte grammatical ou sémantique. Ces résultats sont parfaitement cohérents avec les données cliniques et expérimentales rapportées dans la littérature (Goh *et al.*, 2001; Kaiser *et al.*, 2003; Bergeson *et al.*, 2003, 2005) (voir *« Perception audiovisuelle de la parole après implantation »*, § 9.1.3, p. 193).

La fenêtre temporelle de récupération pour la parole audiovisuelle est beaucoup plus étroite que dans le cas de la parole audiovisuelle, puisque les performances moyennes dans notre groupe de patients implantés se stabilisent dès le deuxième mois suivant la mise en marche de l'implant. Cette rapidité de la récupération audiovisuelle suggère fortement une contribution essentielle de la coopération perceptive entre les voies auditives et visuelles lors de la compréhension audiovisuelle de la parole.

En effet le signal auditif délivré par l'implant cochléaire est considérablement appauvri au niveau spectral, ce qui induit des dégradations de la structure spectro-temporelle fine des signaux sonores. Ces dégradations ont pour conséquence directe une compréhension auditive malaisée de la parole dans des situations comportant un bruit ambiant. Les patients implantés cochléaires ont donc développé des stratégies d'intégration audiovisuelle de la parole spécifiques, qui leur permettent de compenser en grande partie ce désavantage auditif. On pourra se référer en particulier à *« Intégration audiovisuelle chez le sujet implanté »*, sect. 12.2, p. 262, pour une investigation approfondie portant sur ces mécanismes audiovisuels d'intégration.

12.2 Intégration audiovisuelle chez le sujet implanté

12.2.1 Matériel et méthodes

Dans cette étude, nous comparons l'intégration audiovisuelle de la parole chez les patients implantés cochléaires et les sujets normo-entendants au travers de deux modèles mathématiques originaux permettant de simuler une intégration minimale et une intégration optimale. Cette analyse regroupe les données en provenance de quatre études :

- notre étude rétrospective portant sur la perception de la parole auditive, visuelle et audiovisuelle dans un groupe de 97 patients implantés cochléaires (*« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258) ;
- notre étude prospective portant sur la perception de la lecture labiale dans un groupe de 42 sujets normo-entendants (*« Lecture labiale chez le sujet sain »*, sect. 11.1, p. 232) ;
- notre étude prospective sur la perception de la parole auditive et audiovisuelle après dégradation par un masque de bruit blanc dans un groupe de 80 sujets normo-entendants (*« Sujet sain : parole masquée par du bruit »*, sect. 11.2, p. 235) ;
- notre étude prospective sur la perception de la parole auditive et audiovisuelle après simulation acoustique du traitement par l'implant cochléaire dans un groupe de 41 sujets normo-entendants (*« Sujet sain : parole traitée par l'implant »*, sect. 11.3, p. 242).

Concernant l'étude rétrospective, nous nous sommes restreints au sous-groupe de patients qui avaient été testés jusqu'à au moins un an après implantation, ce qui aboutit à une population de 72 patients implantés.

Modèles d'intégration audiovisuelle

Afin de permettre une évaluation quantitative du degré d'intégration audiovisuelle de la parole chez les patients implantés cochléaires aussi bien que chez les sujets contrôles, nous avons

défini deux modèles d'intégration multisensorielle[1]. Tout d'abord, il est important de noter qu'une augmentation de la performance audiovisuelle n'implique pas nécessairement une meilleure intégration audiovisuelle de la parole ; en effet, une telle augmentation peut dans certains cas être simplement liée à une amélioration perceptive en vision seule ou en audition seule, sans faire appel à un renforcement du couplage perceptif entre audition et vision.

Du point de vue probabiliste, le fait de recevoir des informations linguistiques en provenance de deux modalités perceptives améliore automatiquement la probabilité de reconnaissance, puisqu'un mot peut être identifié aussi bien sur la base des informations visuelles que sur la base des informations auditives. Dans notre premier modèle (*1 : intégration minimale*), nous supposerons que les sujets reconnaissent un mot à partir de ce qu'ils entendent ou de ce qu'ils voient, les traitements des informations auditives et visuelles étant supposés indépendants. Dans ce cadre théorique, un mot ne sera pas reconnu si et seulement si aucun des deux traitements auditif ou visuel n'aboutit à une reconnaissance.

Ainsi, en appelant P_{A} la probabilité de reconnaître un mot à partir des seules informations auditives et P_{V} la probabilité de reconnaître un mot à partir des seules informations visuelles, alors la probabilité de ne pas reconnaître ce mot en présence des informations auditives et visuelles est $(1 - P_{\mathrm{A}})(1 - P_{\mathrm{V}})$. La probabilité théorique de reconnaître ce mot en présence des deux modalités sensorielles est donc finalement, dans le cas d'une intégration minimale :

$$\hat{P}_{\mathrm{AV}}^{(1)} = P_{\mathrm{A}} + P_{\mathrm{V}} - P_{\mathrm{A}}P_{\mathrm{V}}$$

Dans notre second modèle (*2 : intégration optimale*), nous utilisons une approche simplifiée de l'intégration audiovisuelle en supposant qu'un mot est reconnu lorsqu'un nombre suffisant d'indices perceptifs spécifiques à ce mot est reconnu, aussi bien à partir des informations visuelles que des informations auditives. Nous introduisons donc un seuil perceptif T, qui définit le nombre d'indices perceptifs minimal nécessaire à la reconnaissance d'un mot. Ces indices perceptifs peuvent être des signaux spatio-dynamiques particuliers en provenance de la vision (mouvements des lèvres, de la langue, du visage, etc.) ou des signaux spectro-temporels particuliers en provenance de l'audition (traits formantiels, répartition spectrale, etc.). Nous supposons de plus que la qualité des afférences sensorielles contrôle le nombre moyen d'indices perceptifs qui peut être détecté.

La détection des indices perceptifs peut être définie de façon probabiliste : nous supposerons que chacun des indices est detecté avec une probabilité p indépendemment des autres indices. Le nombre moyen d'indices détectés dans un mot est donc $\lambda = Np$, où N est le nombre total d'indices contenu dans le mot et p la probabilité de détecter un indice particulier. Cette probabilité p dépend de la qualité perceptive du signal sensoriel et contrôle le niveau de performance prédit par le modèle. Si N est suffisamment important et p suffisamment faible, le nombre n d'indices détectés à chaque essai suit approximativement une loi de Poisson, et la probabilité de détecter k indices est : $\Pi_\lambda(n = k) = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$. Par conséquent la probabilité de reconnaître le mot, qui correspond à la probabilité que le nombre d'indices perceptifs identifiés n dépasse le seuil perceptif T, s'écrit : $P(\lambda, T) = \Pi_\lambda(n > T)$.

Ce formalisme permet de calculer le paramètre λ de la loi de Poisson à partir du taux de reconnaissance de la parole P (indiqué par les résultats expérimentaux) et du seuil perceptif T, qui est un paramètre libre du modèle dont la valeur doit être estimée. Plus cette valeur est importante, plus le score perceptif de l'intégration multisensorielle est proche du score unimodal.

[1] Ces modèles sont dûs à l'amabilité de Sophie Denève, du Groupe de neurosciences théoriques de l'Ecole normale supérieure d'Ulm

En cherchant à minimiser la variance entre les prédictions de notre modèle et les performances individuelles des sujets, nous avons retenu une valeur de seuil $T = 6$, correspondant à une erreur minimale entre les performances théoriques et les performances expérimentales pour les patients implantés et les sujets normaux en condition de masquage par du bruit blanc. Il est intéressant de noter que le seuil perceptif $T = 6$ ainsi obtenu est tout à fait comparable au nombre moyen de phonèmes rencontrés dans un mot disyllabique.

Il est à présent possible de déduire, à partir des scores perceptifs obtenus dans les différentes conditions expérimentales A, V et AV, les paramètres des lois de Poisson correspondantes λ_{A}, λ_{V} et λ_{AV}. Par exemple, λ_{A} est la valeur pour laquelle $P_{\mathrm{A}} = \Pi_{\lambda_{\mathrm{A}}}(n > T)$ est égale à la performance auditive observée expérimentalement. Si les informations auditives et visuelles étaient combinées optimalement, alors il y aurait cumulation des indices perceptifs associés, et le nombre total d'indices perceptifs serait égal à la somme du nombre d'indices perceptifs auditifs et du nombre d'indices perceptifs visuels : $\hat{\lambda}_{\mathrm{AV}} = \lambda_{\mathrm{A}} + \lambda_{\mathrm{V}}$. Nous pouvons donc en déduire la performance théorique qui serait obtenue par un observateur en cas d'intégration optimale, donnée par :

$$\hat{P}^{(2)}_{\mathrm{AV}} = \Pi_{\lambda_{\mathrm{A}}+\lambda_{\mathrm{V}}}(n > T)$$

12.2.2 Résultats et discussion

Les résultats de notre étude rétrospective (cf. *« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258) indiquent que les patients implantés cochléaires ont développé lors de leur surdité des compétences spécifiques en lecture labiale, ce qui leur permet d'atteindre des scores de reconnaissance supra-normaux par comparaison avec les sujets contrôles normo-entendants. Nous avons fait l'hypothèse que cette capacité perceptive développée pouvait être accompagnée d'une meilleure intégration audiovisuelle de la parole, qui permettrait alors de compenser la dégradation des informations auditives par l'implant cochléaire.

Afin de comparer de façon détaillée les capacités d'intégration audiovisuelle des patients implantés avec celles des sujets contrôles normo-entendants, nous avons testé ces derniers en conditions auditive et audiovisuelle dans deux situations d'audition dégradée : 1° dans un premier protocole expérimental (cf. *« Sujet sain : parole masquée par du bruit »*, sect. 11.2, p. 235), nous avons testé 80 sujets en dégradant la composante auditive de la parole par un masque de bruit blanc appliqué à différents rapports signal/bruit ; 2° dans un second protocole expérimental (cf. *« Sujet sain : parole traitée par l'implant »*, sect. 11.3, p. 242), nous avons testé 41 sujets en dégradant la composante auditive de la parole par une simulation acoustique (« vocoder ») du traitement effectué par l'implant cochléaire avec différentes résolutions spectrales.

Dans les deux groupes de sujets, la présentation de parole audiovisuelle induit une amélioration perceptive par rapport à la présentation de parole uniquement auditive. Toutefois, lorsque l'on compare les sujets normo-entendants aux patients implantés cochléaires pour un même niveau de performance auditive, on remarque que les patients implantés obtiennent de meilleures performances audiovisuelles (cf. fig. 12.2, p. 265), cette différence étant très marquée lorsque les sujets contrôles sont exposés à une simulation acoustique d'implant cochléaire.

Cette différence peut recevoir deux explications distinctes mais compatibles : le niveau supranormal de reconnaissance audiovisuelle de la parole chez les patients implantés cochléaires pourrait être lié à leurs performances visuelles supranormales, mais également à un meilleur couplage des informations visuelles avec les informations auditives. La comparaison des gains audiovisuels montre que les patients implantés réalisent un meilleur gain audiovisuel que les sujets normaux, quelle que soit la condition de dégradation de l'audition (cf. fig. 12.3, p. 266). Les prédictions de nos modèles d'intégration minimale et optimale tendent à privilégier la seconde hypothèse, en indiquant une intégration de la parole audiovisuelle sub-optimale chez les sujets

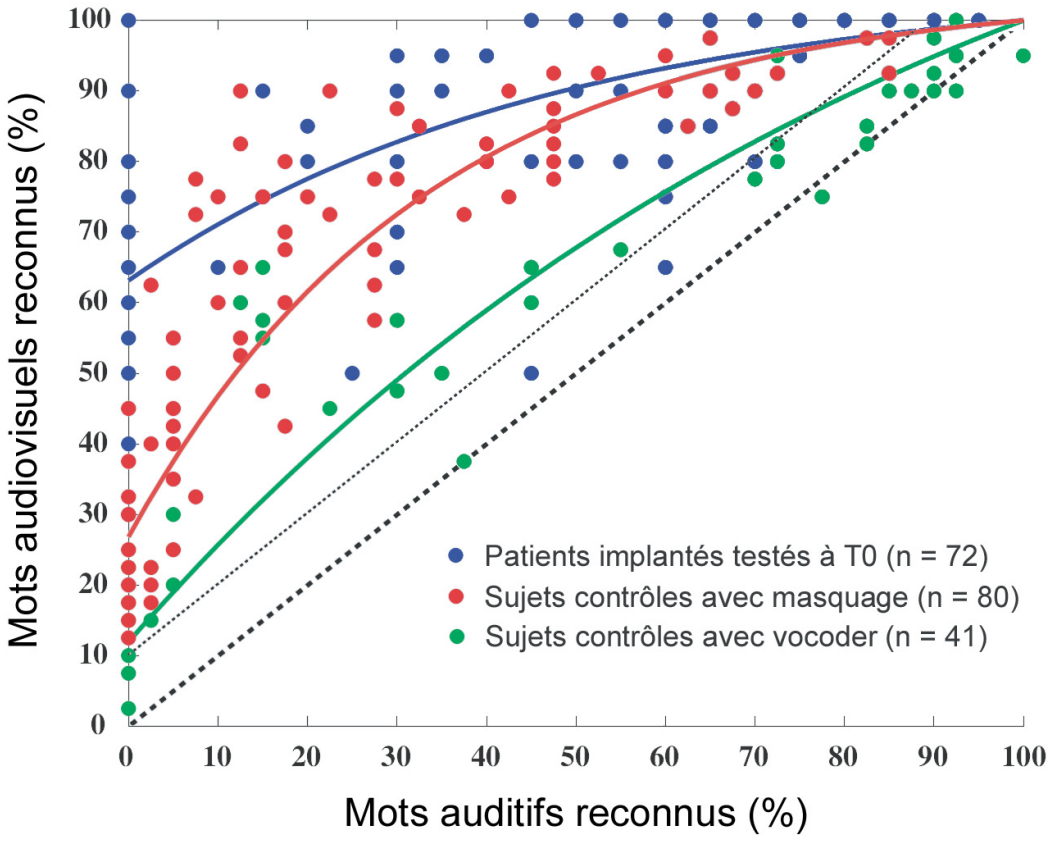


FIG. 12.2 – Comparaison des résultats expérimentaux de reconnaissance audiovisuelle de la parole entre le groupe de sujets contrôles et le groupe de patients implantés. Les niveaux de reconnaissance de la parole en condition audiovisuelle sont comparés pour un même niveau de reconnaissance auditive.

normo-entendants exposés à de la parole vocodée, tandis que cette intégration audiovisuelle devient optimale chez les patients implantés, dès les premiers jours suivant la mise en marche de l'implant (cf. fig. 12.4, p. 267).

Nous avons analysé de façon plus fine cette différence d'intégration audiovisuelle en comparant dans un premier temps les sous-groupes de sujets contrôles et de patients implantés ne parvenant pas à comprendre les mots présentés en audition seule. Cette analyse nous a permis d'isoler des sous-groupes de 13 patients implantés et 19 sujets contrôles, tous deux à 0% de reconnaissance auditive. Les performances audiovisuelles correspondantes sont de 63.0% pour le groupe de patients implantés, tandis qu'elle atteignent 25.4% et 12.5% pour les sujets contrôles exposés à de la parole respectivement masquée et vocodée.

Dans un second temps, nous avons examiné la présence d'une éventuelle corrélation entre les performances visuelles et audiovisuelles des patients implantés afin de tester l'hypothèse selon laquelle l'intégration audiovisuelle supranormale observée serait uniquement liée à leurs performances visuelles supranormales. Bien qu'aucune corrélation ne soit observée ($p > 0.05$) entre les performances visuelles et le gain audiovisuel des patients implantés, nous avons affiné l'analyse en comparant les gains audiovisuels pour un même niveau de lecture labiale dans les deux groupes de sujets. Nous avons donc isolé un sous-groupe de patients implantés ne reconnaissant pas plus de 20% des mots en condition visuelle ($n = 15$). Dans ce groupe, le gain audiovisuel moyen observé

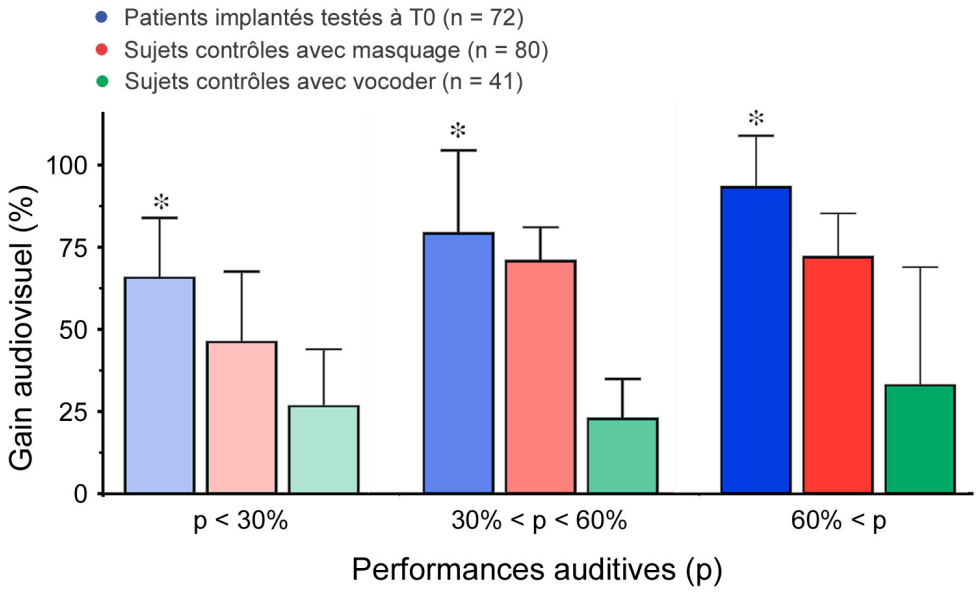


FIG. 12.3 – Comparaison des résultats expérimentaux de reconnaissance audiovisuelle de la parole entre le groupe de sujets contrôles et le groupe de patients implantés. Les gains audiovisuels sont comparés pour un même niveau de reconnaissance auditive. Les astérisques indiquent une différence significative entre les gains audiovisuels des patients implantés et ceux des sujets contrôles.

était de 52%, contre 26% dans le groupe de sujets contrôles, la différence étant statistiquement significative ($p < 0.001$). Cette observation renforce notre hypothèse d'une meilleure intégration audiovisuelle développée par les patients implantés cochléaires.

12.2.3 Conclusion

La comparaison des performances comportementales des patients implantés et des sujets normo-entendants lors de la perception audiovisuelle de la parole démontre que pour des niveaux comparables de performances unimodales visuelles ou auditives, les patients implantés aboutissent à de meilleures performances audiovisuelles, caractérisées à la fois par un meilleur gain audiovisuel et par une intégration audiovisuelle optimale. L'analyse longitudinale des performances montre également que cette optimalité de l'intégration audiovisuelle est atteinte très rapidement par les patients implantés, dès le premier mois suivant la mise en marche de l'implant cochléaire.

Ces résultats mettent donc en évidence le développement rapide d'une coopération forte entre les modalités auditive et visuelle lors de la perception audiovisuelle de la parole chez les patients implantés cochléaires, cette coopération permettant d'atteindre un niveau optimal d'intégration des informations audiovisuelles et de compenser la dégradation des informations auditives par le processeur de l'implant cochléaire. Ces observations suggèrent fortement l'intervention de mécanismes plastiques de réorganisation corticale ; cette problématique particulière sera abordée dans notre étude longitudinale en neuroimagerie cérébrale (« *Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral* », ch. 13, p. 289).

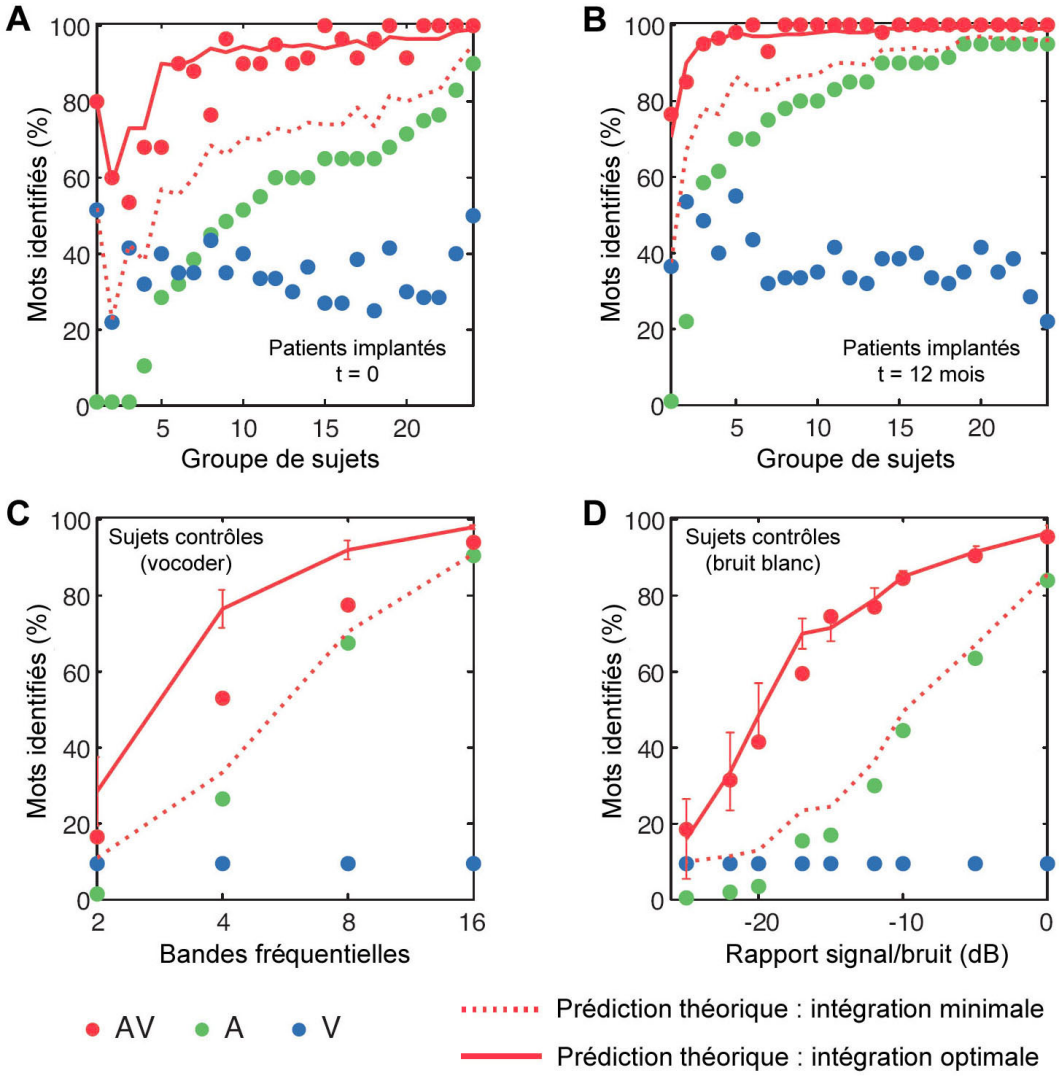


FIG. 12.4 – Comparaison des résultats expérimentaux de reconnaissance audiovisuelle de la parole avec les prédictions des modèles 1 (intégration minimale) et 2 (intégration optimale). A et B : patients implantés testés à la mise en marche (A) puis un an après la mise en marche (B) de l'implant cochléaire (n = 72) ; C : sujets contrôles testés avec une simulation acoustique de parole traitée par l'implant cochléaire (n = 41) ; D : sujets contrôles testés avec de la parole dégradée par un masque de bruit blanc (n = 80). Pour C et D, le niveau d'identification des mots en condition visuelle (lecture labiale) est fixé à 9.4% (n = 42). Tous les sujets ont été classés par ordre de performance auditive. Pour faciliter la lisibilité des graphiques, les patients ont été regroupés par sous-groupes de 3 sujets.

12.3 Lecture labiale chez le sujet implanté

12.3.1 Matériel et méthodes

Sujets

Dans cette étude nous effectuons une comparaison approfondie entre les performances comportementales de lecture labiale obtenues dans un groupe de 97 patients implantés testés lors de notre étude rétrospective longitudinale (*« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258) et celles obtenues dans un groupe de 42 sujets normo-entendants testés dans notre étude prospective (*« Lecture labiale chez le sujet sain »*, sect. 11.1, p. 232). Afin de permettre une comparaison longitudinale, nous avons divisé le groupe de patients implantés en trois sous-groupes :

- un sous-groupe de patients sourds postlinguaux testés au plus 18 mois avant implantation cochléaire (41 femmes et 24 hommes) ;
- un sous-groupe de patients implantés testés entre 3 et 6 mois après implantation cochléaire (49 femmes et 31 hommes) ;
- un sous-groupe de patients implantés testés entre 24 et 36 mois après implantation cochléaire (28 femmes et 18 hommes).

En raison des modalités de déroulement du suivi orthophonique, seule une moitié des patients du groupe clinique ont subi des tests pour les trois périodes que nous avons isolées (pré-implantation, post-implantation précoce et post-implantation tardive), tandis que les autres n'ont été testés que pour une ou deux de ces périodes. Cela explique la différence du nombre de sujets en fonction des périodes de test, notre étude ne pouvant pour cette même raison pas être considérée à strictement parler comme une étude longitudinale.

Analyse statistique

L'hypothèse nulle de normalité des performances comportementales de nos patients implantés a été rejetée par plusieurs tests statistiques (Kolmogorov-Smirnov, Lilliefors et Bera-Jarque) pour un seuil de significativité fixé à $p = 0.05$. Comme nos données présentent une déviation par rapport à la loi normale et incluent des mesures répétées, nous avons choisi d'utiliser la méthode du *bootstrap* pour l'analyse statistique (Carpenter et Bithell, 2000). Cette méthode présente également l'avantage de compenser des différences dans la taille des groupes étudiés, en fournissant un nombre fixe de simulations (ici, $n = 10\,000$) par groupe pour chaque condition. Tous les calculs ont été réalisés sous l'environnement statistique R (R, 2007), en utilisant la librairie *boot*.

Les données de chaque sous-groupe (hommes et femmes) pour chaque condition (groupe normo-entendant, groupe pré-implantation, groupe post-implantation précoce et groupe post-implantation tardive) ont été rééchantillonnées 10 000 fois, aboutissant à chaque fois à une moyenne d'échantillon. Nous avons donc obtenu une distribution de 10 000 observations simulées pour chaque sous-groupe et chaque condition. Nous avons ensuite utilisé des intervalles de confiance accélérés et centrés (Carpenter et Bithell, 2000) pour un seuil de significativité égal à 0.05 afin d'estimer la significativité des différences entre les réponses des hommes et des femmes.

Les gains perceptifs des femmes par rapport aux hommes ont été estimés selon la formule $\frac{F_m - H_m}{F_m}$, où H_m et F_m sont les vecteurs obtenus par 25 rééchantillonages des réponses comportementales obtenues par les hommes et les femmes. La même procédure de rééchantillonage que pour les scores globaux a ensuite été appliquée afin d'estimer la significativité des différences de gain entre les hommes et les femmes.

12.3.2 Résultats et discussion

Performances globales

Les intervalles de confiance au seuil $p = 0.05$ obtenus par la méthode du *bootstrap* sont respectivement de $9.4 \pm 1.7\%$ pour les sujets normo-entendants, $29.7 \pm 3.5\%$ pour les patients avant implantation, $36.3 \pm 2.6\%$ pour les patients testés précocément après implantation et $37.7 \pm 3.5\%$ pour les patients testés tardivement après implantation. Par conséquent, en établissant la significativité des différences sur la base du non-chevauchement des intervalles de confiance, il n'y a aucune différence significative entre les patients testés précocément ou tardivement après implantation cochléaire. Les patients testés après implantation présentent une amélioration significative de leurs performances en lecture labiale par rapport aux patients testés avant implantation, ces derniers étant à leur tour plus performants que les sujets normo-entendants.

Différences hommes/femmes

Lorsque les sous-groupes d'hommes et de femmes sont analysés séparément, on trouve les intervalles de confiance suivants :

- $12.3 \pm 2.8\%$ pour les femmes et $6.6 \pm 1.3\%$ pour les hommes chez les sujets normo-entendants (gain des femmes : $44.5 \pm 0.5\%$) ;
- $33.3 \pm 4.1\%$ pour les femmes et $23.5 \pm 5.2\%$ pour les hommes chez les patients testés avant implantation (gain des femmes : $29.2 \pm 2.5\%$) ;
- $39.5 \pm 3.2\%$ pour les femmes et $31.3 \pm 3.7\%$ pour les hommes chez les patients testés précocément après implantation (gain des femmes : $21.0 \pm 2.6\%$) ;
- $37.9 \pm 5.5\%$ pour les femmes et $37.5 \pm 4.7\%$ pour les hommes chez les patients testés tardivement après implantation (gain des femmes : $2.5 \pm 2.5\%$).

Par conséquent, l'avantage perceptif des femmes en lecture labiale est significatif pour toutes les populations, à l'exception des patients testés tardivement après implantation cochléaire. Cette différence hommes/femmes est la plus importante chez les sujets normo-entendants (44.5%), suivis par les patients sourds avant implantation cochléaire (29.2%), puis par les patients implantés testés précocément (21.0%), puis enfin par les patients implantés testés tardivement, pour lesquels la différence hommes/femmes n'est plus significative.

Discussion

Nos résultats montrent l'existence d'une différence hommes/femmes lors de la perception de la parole visuelle, les femmes étant nettement plus avantagées que les hommes lors d'une tâche de lecture labiale. Nos données cliniques montrent également que l'avantage perceptif des femmes par rapport aux hommes persiste en cas de surdité, avec toutefois une diminution de l'écart hommes/femmes. En cas d'implantation cochléaire, cet écart diminue encore et on assiste à un « rattrapage » des femmes par les hommes sur le long terme ; en effet la différence hommes/femmes est encore significative pour des patients testés dans le premier trimestre après l'implantation, mais elle se réduit jusqu'à devenir non-significative plus de deux ans après implantation cochléaire.

Ces résultats confirment et étendent les données comportementales et neurofonctionnelles disponibles dans la littérature concernant les différences hommes/femmes lors de la lecture labiale. En effet un ensemble d'études comportementales indique que les femmes obtiennent des performances de lecture labiale significativement plus importantes que les hommes (Johnson et Leslie, 1988; Watson *et al.*, 1996). Par ailleurs certaines études démontrent une plus grande influence de la composante visuelle de la parole chez les femmes lors de la perception de stimuli de

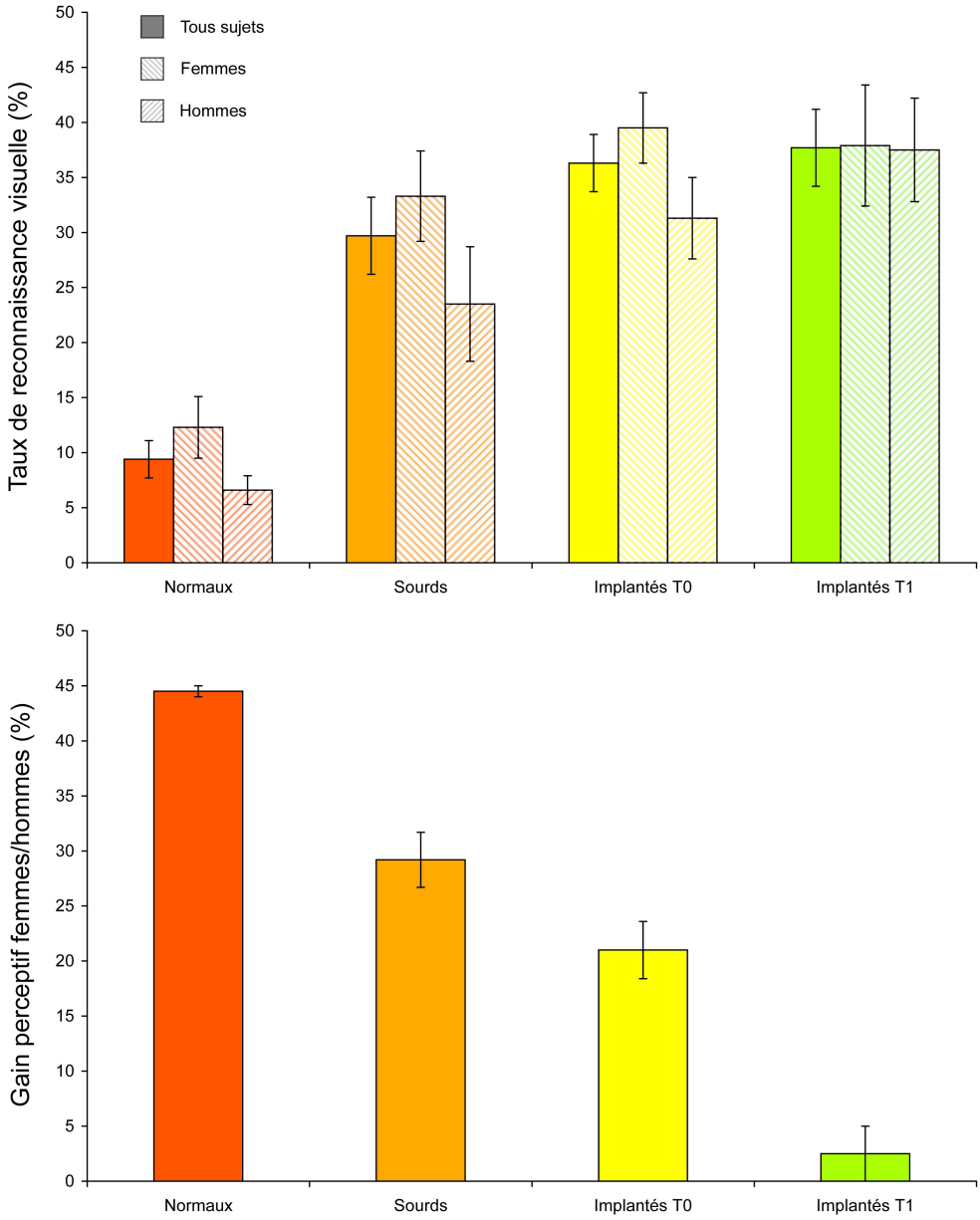


FIG. 12.5 – En haut : performances de reconnaissance des mots en condition de parole visuelle (%) ; en bas : gain perceptif des femmes par rapport aux hommes (%). Observation dans un groupe de sujets normo-entendants (20F/22H : *Normaux*), un groupe de patients sourds avant implantation (41F/24H : *Sourds*), un groupe de patients sourds testés dans le premier trimestre après implantation cochléaire (49F/31H : *Implantés T0*) et un groupe de patients sourds testés plus de deux ans après implantation cochléaire (28F/18H : *Implantés T1*). Les barres d'erreur indiquent les intervalles de confiance obtenus par la méthode du *bootstrap* pour un seuil p = 0.05 ; des intervalles se chevauchant indiquent donc une différence non significative.

parole audiovisuelle incongruente de type McGurk (Aloufy *et al.*, 1996; Irwin *et al.*, 2006) (voir *« Différence d'intégration audiovisuelle entre hommes et femmes »*, § 3.6.4, p. 71 et *« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91). L'avantage perceptif des femmes pourrait

être expliqué dans le cadre du modèle d'activation-vérification proposé par Strelnikov (2007), les femmes utilisant des stratégies intégratives et prédictrices de façon plus extensive que les hommes.

La diminution progressive de l'écart hommes/femmes entre la période de surdité, les premiers mois après implantation et la période de long terme après implantation peut être expliquée par la modification progressive des stratégies cognitives utilisées par les hommes. Dans ce cadre il y aurait donc une transition depuis la période d'audition normale, où les hommes font un usage restreint des stratégies intégratives et prédictrices tandis que les femmes les utilisent régulièrement, vers la période de surdité puis d'implantation cochléaire, qui induisent des demandes croissantes en stratégies compensatrices. Les hommes, dont les stratégies perceptives lors de la compréhension du langage ne sont pas optimales, bénéficieraient ainsi d'une plus grande marge de manœuvre que les femmes, dont les stratégies perceptives sont plus proches de l'optimalité, ce qui expliquerait l'importante marge de progression des hommes.

12.4 Effet McGurk chez le sujet implanté

12.4.1 Matériel et méthodes

Sujets

Nous avons réalisé une étude prospective sur un groupe de 33 patients adultes présentant une surdité profonde bilatérale postlinguale et ayant subi une implantation cochléaire unilatérale, ainsi que sur un groupe témoin de 39 sujets adultes ne présentant aucune déficience auditive. Le groupe de sujets implantés est composé de 16 femmes et 17 hommes entre 19 et 76 ans, avec une moyenne d'âge de 52.0 ans (± 15.8 ans). Le groupe de sujets sains est composé de 20 femmes et 19 hommes entre 19 et 58 ans, avec une moyenne d'âge de 32.0 ans (± 13.0 ans). Les étiologies des patients sourds implantés sont très variables, principalement infectieuses et génétiques (méningite, otite chronique, otosclérose, neurinome acoustique, otospongiose, etc.) Tous les patients ont reçu un implant cochléaire Nucleus (CI-22 ou CI-24), avec l'une des stratégies de codage MPEAK, SPEAK ou ACE. Cette investigation a été effectuée dans le service ORL du CHU Toulouse-Purpan pour les sujets implantés et dans le centre Cerveau et Cognition UMR CNRS 5549 à la faculté de médecine de Toulouse-Rangueil pour les sujets sains. Tous les sujets étaient entièrement naïfs vis-à-vis du protocole expérimental, aucun d'entre eux n'ayant jamais été exposé auparavant aux stimuli utilisés.

Stimuli

Préparation des stimuli : Les stimuli sont tous des mots dissyllabiques de type VCV, où V est systématiquement la voyelle /a/, et où C est une des huit consonnes occlusives /p/, /b/, /m/, /t/, /d/, /n/, /k/, /g/ ou l'une des six consonnes constrictives /f/, /v/, /ʃ/, /ʒ/, /s/, /z/. L'enregistrement des prononciations de chaque stimulus s'est déroulé suivant des modalités similaires à celles décrites dans *« Génération des stimuli »*, § 10.2.1, p. 218, l'orthophoniste ayant effectué trois répétitions de chacun des quatorze mots dissyllabiques.

Le doublage des stimuli de type McGurk a été effectué sur un processeur Pentium 4 2.4 GHz sous Adobe Premiere Pro 7.0. Six doublages ont été effectués : les trois stimuli auditifs /apa/, /aba/ et /ama/ (occlusives labiales) ont été doublés par le stimulus visuel /aka/ (occlusive vélaire sourde), puis les trois stimuli auditifs /ata/, /ada/ et /ana/ (occlusives dentales) ont été doublés par le stimulus visuel /apa/ (occlusive labiale sourde). Le doublage a été systématiquement effectué en assurant pour chaque mot une parfaite coïncidence temporelle entre les transitions

V-C du son original et du son de doublage, ainsi qu'une coïncidence temporelle optimale entre les temps d'émission ou de relâchement des voyelles.

Les stimuli ont été montés ensemble sous forme de films vidéo au format MPEG2 avec Adobe Premiere Pro 7.0. Quatre conditions ont été testées : une condition visuelle seule (V), pour laquelle seule l'image était présentée, le son étant coupé ; une condition auditive seule (A), pour laquelle seul le son était présenté, l'écran restant totalement noir ; une condition audiovisuelle congruente (AV) pour laquelle l'image et le son de l'enregistrement original étaient présentés conjointement ; enfin une condition audiovisuelle non-congruente (AVMG) pour laquelle un doublage de type McGurk était présenté.

Ordre de présentation : Pour l'expérience I, la condition auditive incluait les huit consonnes occlusives ainsi que les deux constrictives /f/ et /v/, la condition visuelle incluait les huit consonnes occlusives, la condition audiovisuelle congruente incluait les huit consonnes occlusives et la condition audiovisuelle non-congruente incluait les trois stimuli de type McGurk correspondant aux stimuli auditifs labiaux /apa/, /aba/ et /ama/ doublés par le stimulus visuel /aka/.

Pour l'expérience II, les conditions auditive seule et visuelle seule incluaient les huit consonnes occlusives et les six consonnes constrictives. La condition audiovisuelle congruente incluait les huit consonnes occlusives et la condition audiovisuelle non-congruente incluait les six stimuli de type McGurk (trois stimuli auditifs labiaux et trois stimuli auditifs dentaux).

Dans chaque expérience, les stimuli ont été regroupés en petits blocs correspondant à une même condition perceptive : auditive seule, visuelle seule et audiovisuelle (cette condition incluant les stimuli congruents et les stimuli de type McGurk). Les blocs comprenaient 2 à 3 stimuli pour l'expérience I et 4 à 5 stimuli pour l'expérience II. Tous les blocs étaient présentés dans un ordre aléatoire pour chaque sujet, en respectant toutefois une alternance des différentes conditions perceptives (A, V et AV). À l'intérieur de chaque bloc, les stimuli étaient choisis aléatoirement pour chaque sujet.

Protocole

Les stimuli ont été présentés sur un ordinateur portable avec un moniteur de dimensions 20.7×33.1 cm. L'image était présentée au centre de l'écran avec une dimension fixe de 20.7×27.6 cm, le reste de l'écran demeurant noir. Les sujets étaient assis confortablement en face de l'écran, à une distance de lecture normale d'environ 65 cm, ce qui correspond à un diamètre angulaire d'environ 30°. Le son était délivré à travers les haut-parleurs placés de chaque côté du moniteur, à un niveau sonore de 65 dB et avec une faible déformation spectrale.

Les participants ont tous été testés individuellement. Ils ont reçu pour seules instructions de regarder l'image et d'écouter le son pendant toute la durée de présentation de chaque stimulus, puis de répéter à voix haute ce qu'ils avaient perçu. L'expérimentateur était assis dans la pièce pendant toute la durée de l'expérience pour confirmer que les sujets regardaient bien l'écran et pour consigner par écrit leurs réponses verbales. Une transcription phonétique systématique a été effectuée pour chacune de ces réponses.

Analyse des résultats

Classification des réponses : Seul le groupe de consonnes central des réponses données par les sujets a été pris en compte pour l'analyse des réponses, indépendemment des voyelles entourant ce groupe central. Dans les cas où un groupe de deux consonnes était rapporté, la règle suivante a été appliquée : 1° si les deux consonnes sont caractérisées par le même mode d'articulation, le mode d'articulation associé à la réponse est le mode commun aux deux consonnes ; 2° si les deux

consonnes sont caractérisées par des modes d'articulation différents, la réponse est considérée comme ambigüe et n'est pas incluse dans l'analyse. En pratique, très peu (moins de 1.3%) de percepts incongruents se sont révélés ambigüs, dans le groupe de patients implantés comme dans le groupe de sujets contrôles.

Dans le cas d'un stimulus audiovisuel incongruent, on peut subdiviser les percepts obtenus en quatre classes distinctes : dominance auditive, dominance visuelle, fusion ou combinaison (pour une explication plus détaillée, voir *« Conséquences d'une incongruence phonétique : effet McGurk »*, sect. 3.5, p. 65). C'est cette classification qui a été utilisée pour caractériser les percepts rapportés par les sujets. Tous les sujets (patients et contrôles) ont spontanément donné une réponse pour tous les stimuli audiovisuels incongruents.

Catégorie	*Stimulus audiovisuel*		**Percept**
	Auditif	**Visuel**	
Dominance auditive	/aba/	/aga/	/**aba**/
Dominance visuelle	/aba/	/aga/	/**aga**/
Fusion	/aba/	/aga/	/**ada**/
Combinaison	/aga/	/aba/	/**abga**/

TAB. 12.1 – Catégories perceptives pour un stimulus audiovisuel incongruent

ANALYSE STATISTIQUE : L'analyse statistique des résultats a été effectuée sous le logiciel StatView. L'hypothèse nulle de normalité des données ayant été rejetée par plusieurs tests statistiques, l'analyse statistique a été effectuée à l'aide de tests non paramétriques. Le test non-apparié de Mann-Whitney a été utilisé pour comparer les réponses des patients et des contrôles pour une même condition expérimentale. Le test apparié de Wilcoxon a été utilisé pour comparer les réponses d'une même population à deux conditions expérimentales différentes. Comme les résultats communs aux deux expériences (I et II) ne montraient aucune différence statistiquement significative, les données des deux expériences ont été regroupées ensemble en vue d'une analyse globale.

ANALYSE PHONÉTIQUE DES CONFUSIONS : Les réponses des sujets ont subi une classification phonétique complète, et les résultats de cette analyse ont été représentés sous forme de matrices de confusions phonétiques, dont le principe est détaillé dans *« Matrices de confusions phonétiques »*, § 10.3.2, p. 225. Nous avons également effectué une analyse de transmission d'information à partir de l'analyse phonétique des réponses des sujets ; les principes de base de cette analyse sont résumés dans *« Analyse de la transmission d'information phonétique »*, § 10.3.3, p. 225.

12.4.2 Résultats

Stimuli congruents

CONDITION AUDITIVE SEULE : dans le cas d'une stimulation uniquement auditive, les sujets contrôles reconnaissent parfaitement tous les stimuli proposés, avec un score auditif de 99.0 ± 3.4% (voir fig. 12.6, p. 274). Ce résultat est également illustré dans les matrices de confusion, par la répartition diagonale des réponses (voir fig. 12.7, p. 275, en haut et à droite).

Les sujets implantés rencontrent au contraire d'importantes difficultés de reconnaissance des phonèmes, avec un score auditif de 29.6 ± 18.7%. Sur les matrices de confusion, cette difficulté

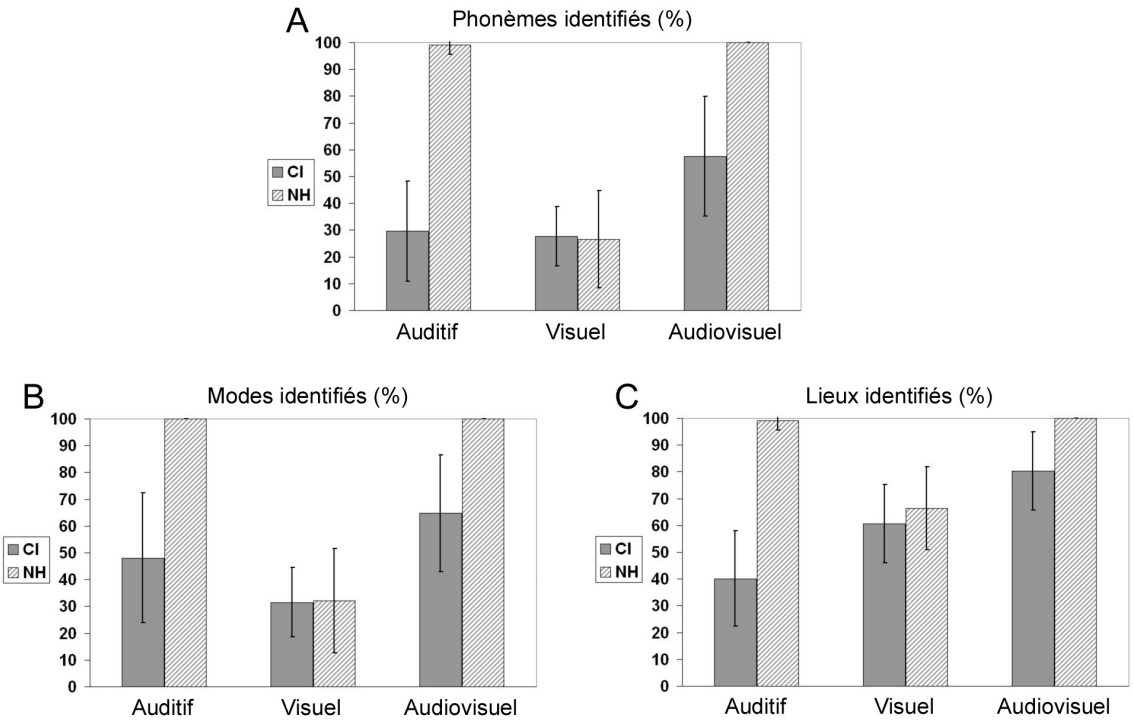


FIG. 12.6 – Stimuli unimodaux et audiovisuels congruents : identifications et catégorisations articulatoires en lieux et modes d'articulation. Observations dans un groupe de 33 patients implantés cochléaires et 39 sujets contrôles.

perceptive se traduit par une dispersion des réponses entre de nombreux percepts possibles avec une faible dominance diagonale (figure 12.7, en haut et à gauche). En particulier, la dominance diagonale semble la moins marquée dans le cas des bilabiales /p/ /b/ /m/ et des dentales /t/ /d/ /n/.

L'analyse des réponses par transmission d'information phonétique montre que la perception auditive des patients implantés ne transmet que 23.1% des informations de voisement et 21.5% des informations de nasalité. Concernant les lieux d'articulation, la transmission auditive de l'information articulatoire chez les patients implantés est de 21.7% pour les vélaires, 5.4% pour les bilabiales et 1.8% pour les dentales.

CONDITION VISUELLE SEULE : pour chacun des stimuli uniquement visuels (lecture labiale), les sujets contrôles et les patients implantés présentent des performances globales parfaitement similaires, respectivement de 26.6 ± 18.2% et 27.7 ± 11.2% (différence non significative, $p = 0.41$).

La répartition détaillée des réponses entre les différents percepts est également similaire (figure 12.7, en bas). Dans les deux cas, on observe une répartition des percepts en blocs diagonaux, ce qui traduit l'existence de groupes phonétiques perceptuellement stables en lecture labiale. On peut noter en particulier le groupe des occlusives antérieures (bilabiales) /p/ /b/ /m/, celui des occlusives postérieures (dentales et vélaires) /t/ /d/ /n/ /k/ /g/, celui des fricatives labiodentales /f/ /v/, celui des fricatives palatales /ʃ/ /ʒ/ et également celui des fricatives alvéolaires /s/ /z/, bien qu'il demeure de nombreuses confusions entre ce dernier et le groupe des occlusives postérieures.

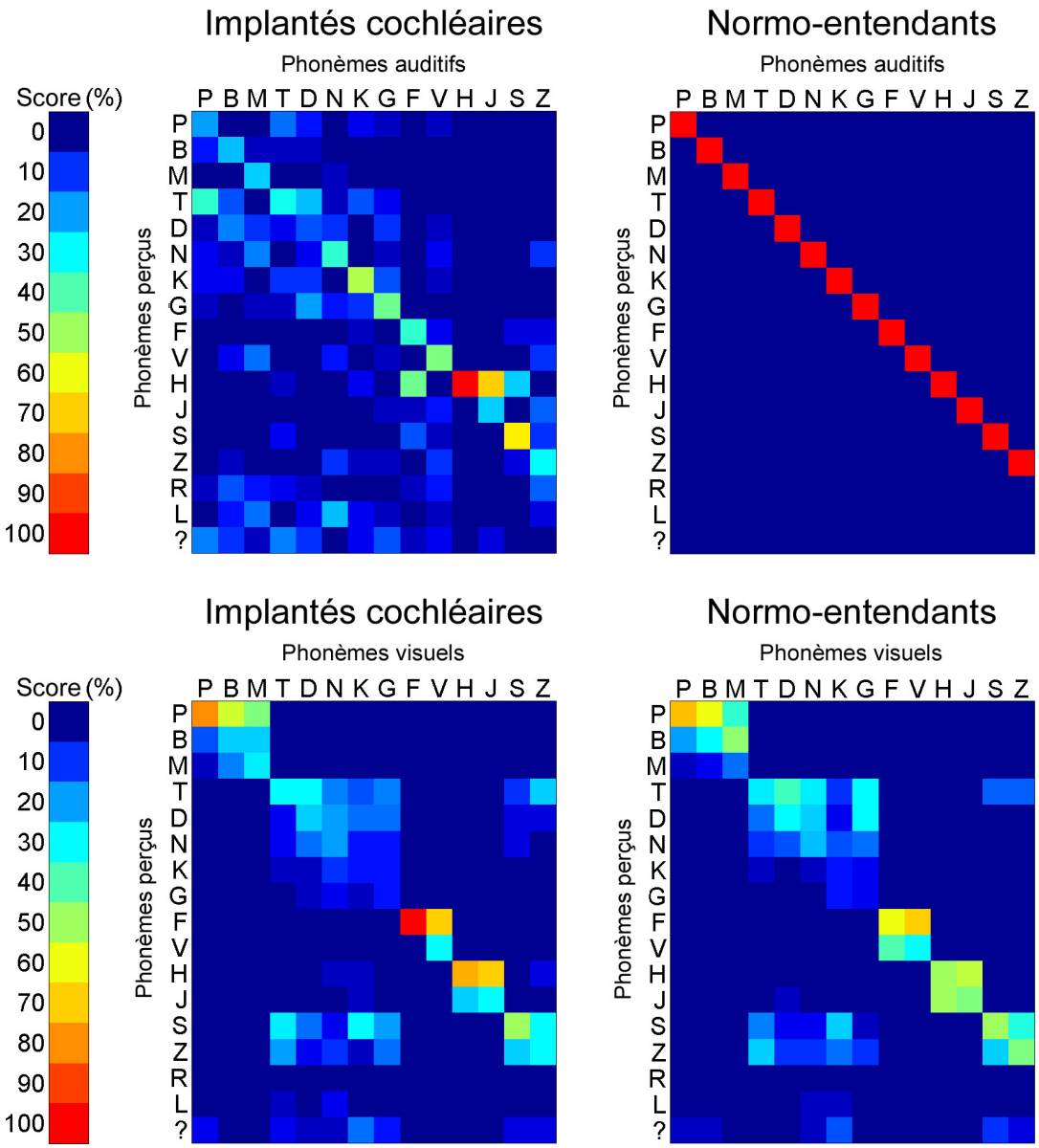


FIG. 12.7 – Matrices de confusion en conditions auditive seule (haut) et visuelle seule (bas).

L'analyse des réponses par transmission d'information phonétique montre que la perception visuelle ne transmet que 3.8% des informations de voisement et 3.7% des informations de nasalité. La transmission visuelle des informations de lieux d'articulation est quant à elle de 96.4% pour les bilabiales, 14.9% pour les dentales et 4.9% pour les vélaires.

CONDITION AUDIOVISUELLE CONGRUENTE : dans le cas d'une stimulation audiovisuelle congruente, les sujets contrôles reconnaissent parfaitement tous les stimuli proposés, avec un score audiovisuel de 100% (voir fig. 12.6, p. 274). Ce résultat est également illustré dans les matrices de confusion par la répartition diagonale des réponses dans les matrices de confusion (voir fig. 12.8, p. 276, partie droite).

Comme dans le cas de la condition auditive seule, les sujets implantés rencontrent des difficultés dans la reconnaissance des phonèmes audiovisuels, avec un score de 57.6 ± 22.3%.

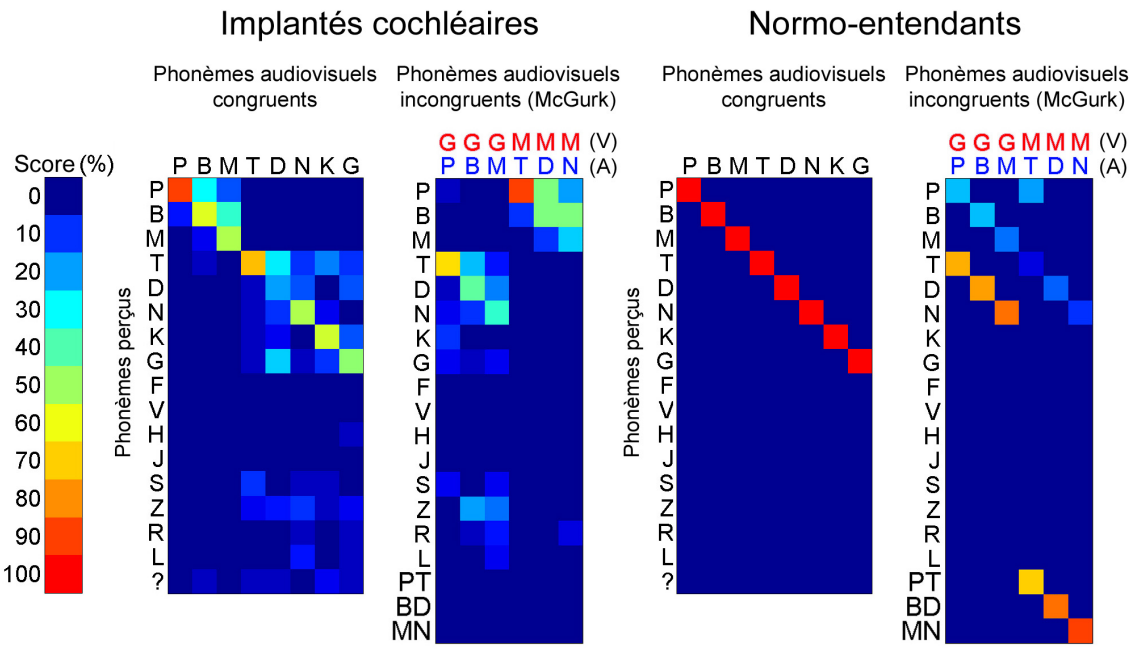


FIG. 12.8 – Matrices de confusion en conditions bimodale congruente et bimodale incongruente.

Toutefois ce score est significativement plus élevé ($p < 0.0001$; Wilcoxon) que dans la condition auditive seule, ce qui montre l'avantage perceptif apporté par la présentation simultanée des informations auditives et visuelles, même dans un contexte uniquement phonétique et dépourvu de tout indice lexical ou sémantique.

L'observation de la répartition détaillée des réponses des sujets dans les matrices de confusion (voir fig. 12.8, p. 276, à gauche) montre l'apparition de deux différences importantes par rapport aux stimuli unimodaux. Tout d'abord, la dominante diagonale est mieux marquée que pour les conditions unimodales auditive et visuelle, ce qui indique l'intervention d'un mécanisme audiovisuel de facilitation perceptive. Ensuite, on peut remarquer une structure en blocs diagonaux identique à celle observée dans le cas visuel unimodal, ce qui montre que l'ajout des informations visuelles aux informations auditives permet d'éliminer des possibilités de confusion en ne conservant que des percepts compatibles avec la composante visuelle de la stimulation.

Stimuli incongruents

PERCEPTION DU VOISEMENT ET DE LA NASALITÉ

En cas d'incongruence audiovisuelle, l'identification du mode de la composante auditive par les sujets normo-entendants est quasi-parfaite (98.3% pour les stimuli McGurk dentaux, et 99.2% pour les stimuli McGurk labiaux), tandis que l'identification du mode de la composante visuelle n'est correcte que dans un cas sur trois (33.3% dans tous les cas), ce qui correpond au taux théorique pour des réponses faites au hasard (voir fig. 12.9, p. 277). L'application de l'analyse par transmission d'information phonétique aux résultats des sujets normaux montre que 99.5% des informations de mode d'articulation sont transmises par l'audition, tandis que seules 0.1% de ces informations sont transmises par la vision.

Comme on peut s'y attendre, les identifications du mode d'articulation sont moins bonnes chez les patients implantés. En cas d'incongruence audiovisuelle, l'identification du mode de la composante auditive se fait de façon correcte pour la moitié des stimuli (55.6% pour les sti-

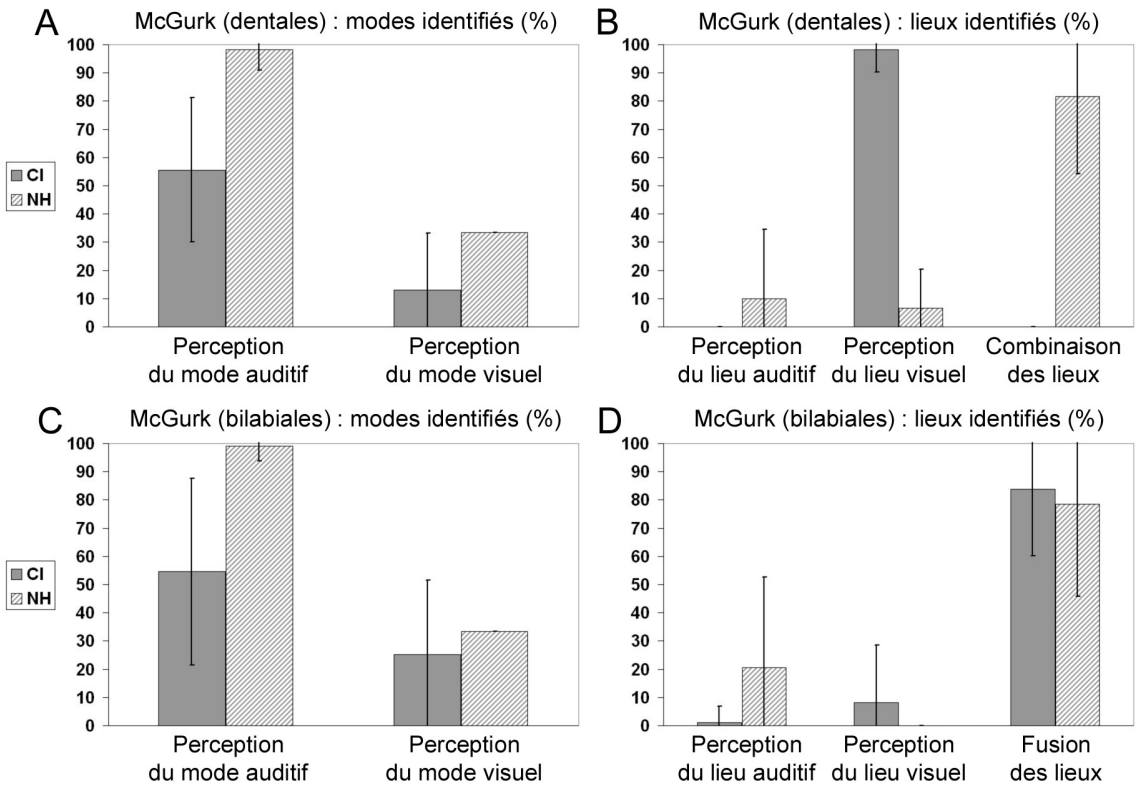


FIG. 12.9 – Stimuli audiovisuels incongruents : catégorisations articulatoires en lieux et modes d'articulation. Observations dans un groupe de 33 patients implantés cochléaires et 39 sujets contrôles.

muli McGurk dentaux, et 54.6% pour les stimuli McGurk labiaux), tandis que l'identification du mode de la composante visuelle n'est correcte que dans moins d'un cas sur quatre (13.0% pour les stimuli McGurk dentaux, et 25.3% pour les stimuli McGurk labiaux), ce qui est inférieur au taux théorique pour des réponses faites au hasard (voir fig. 12.9, p. 277). L'analyse par transmission d'information phonétique montre que pour les patients implantés, 22% des informations de mode d'articulation sont transmises par l'audition, tandis que seules 9% de ces informations sont transmises par la vision.

PERCEPTION DU LIEU D'ARTICULATION

Dans le cas des stimuli McGurk dentaux, c'est-à-dire les stimuli auditifs /ata/, /ada/ et /ana/ présentés avec un /ama/ visuel, la grande majorité (81.7%) des réponses des sujets contrôles sont des combinaisons perceptives du lieu auditif et du lieu visuel (voir fig. 12.9, p. 277), avec les percepts typiques respectifs /apta/, /abda/ et /amna/ (voir fig. 12.8, p. 276). Pour ces mêmes stimuli, une très grande majorité (98.2%) des réponses des patients implantés impliquent uniquement le lieu visuel (voir fig. 12.9, p. 277), avec les percepts typiques respectifs /apa/, /aba/ et /aba/ (voir fig. 12.8, p. 276).

Dans le cas des stimuli McGurk bilabiaux, c'est-à-dire les stimuli auditifs /apa/, /aba/ et /ama/ présentés avec un /aga/ visuel, la grande majorité (78.6%) des réponses des sujets contrôles sont des fusions perceptives du lieu auditif et du lieu visuel (voir fig. 12.9, p. 277), avec les percepts typiques repectifs /ata/, /ada/ et /ana/ (voir fig. 12.8, p. 276). Les patients implantés montrent un comportement similaire, en répondant pour une grande majorité (83.8%)

par les fusions perceptives typiques /ata/, /ada/ et /ana/ (voir fig. 12.8, p. 276).

Ainsi, en cas d'incongruence audiovisuelle, les lieux d'articulations les plus rapportés par les sujets contrôles et les patients implantés ne correspondent pas au lieu d'articulation présenté en audition. Dans le cas des stimuli McGurk dentaux, les lieux auditifs sont rapportés correctement pour 10.0% des réponses chez les sujets contrôles et ne sont jamais rapportés (0%) chez les patients implantés (différence significative ; $p = 0.048$). De même, dans le cas des stimuli McGurk bilabiaux, les lieux auditifs sont rapportés correctement pour 20.5% des réponses chez les sujets contrôles et pour uniquement 1.0% des réponses chez les patients implantés (différence significative ; $p = 0.0006$).

12.4.3 Discussion

Perception unimodale de la parole

Complémentarité de l'audition et de la vision.

Les données de la littérature scientifique indiquent une différence entre les spécialisations phonétiques de l'audition et de la vision pour la perception de la parole. De façon générale, la vision transmet de façon plus pertinente les informations relatives au lieu d'articulation, en particulier pour les lieux antérieurs. D'un autre côté, l'audition transmet mieux les informations portant sur le mode d'articulation, en particulier le voisement et la nasalité, tandis que les informations portant sur le lieu d'articulation sont parmi les premières à subir l'effet d'une dégradation des informations auditives. On pourra se référer à *« Congruence phonétique de la parole : complémentarité audiovisuelle »*, § 3.3.2, p. 63 pour une revue de la littérature à ce sujet.

Dans notre étude, l'analyse phonétique détaillée des percepts unimodaux pour les sujets contrôles et les patients implantés cochléaires se replace complètement dans ce cadre. Ainsi, dans notre groupe de patients, l'implant cochléaire transmet le mieux les informations de voisement (23.1%) et de nasalité (21.5%), comme le confirme l'analyse par transmission d'information phonétique. Parmi les informations de lieu d'articulation, l'implant cochléaire transmet mieux l'information phonétique pour les lieux postérieurs (21.7% pour les vélaires) que pour les lieux antérieurs et intermédiaires, dont l'information est particulièrement dégradée (5.4% pour les bilabiales et 1.8% pour les dentales).

Au niveau de la vision, les informations de voisement et de nasalité sont les plus dégradées (3.8% et 3.7%), alors que ce sont celles qui étaient le mieux transmises par l'implant cochléaire. Parmi les informations de lieu d'articulation, l'information pour les lieux antérieurs est transmise avec une très grande fiabilité par la vision (96.4% pour les bilabiales) alors qu'elle était très dégradée par l'audition. Inversement, l'information pour les lieux postérieurs est la moins bien transmise par la vision (4.7% pour les vélaires) alors qu'elle était la mieux transmise par l'audition.

Le fait que la complémentarité phonétique entre audition et vision soit préservée dans un cas aussi particulier que celui des patients implantés cochléaires confirme que cette propriété est un principe fondamental de la parole audiovisuelle. Nous utiliserons dans la suite ce principe comme postulat de base pour discuter des mécanismes d'intégration audiovisuelle de la parole employés par les patients implantés cochléaires.

Lecture labiale en contexte non linguistique.

Un résultat surprenant au premier abord est la grande similarité des réponses données en condition visuelle seule (lecture labiale) entre les sujets contrôles et les patients implantés. Cette similarité se retrouve à tous les niveaux d'analyse : en performance globale, en catégorisation phonétique détaillée ainsi qu'en analyse par transmission d'information.

Ce résultat peut paraître contradictoire avec les compétences de lecture labiale spécifiques qu'ont développé les patients implantés cochléaires (voir *« Perception visuelle de la parole après implantation »*, § 9.1.2, p. 193 ; voir également notre étude précédente *« Lecture labiale chez le sujet implanté »*, sect. 12.3, p. 268). Toutefois il faut prendre en compte la difficulté supplémentaire essentielle introduite par notre protocole expérimental, qui fait intervenir des associations de phonèmes sans signification. Cette particularité a pour effet de supprimer tout indice de nature linguistique, qu'il soit lexical, sémantique ou syntaxique.

Nos données expérimentales mettent donc en valeur la nature circonstancielle des capacités de lecture labiale présentes chez les patients implantés. Elles tendent à indiquer que les compétences perceptives visuelles des patients implantés ne s'exercent pas à un bas niveau de traitement cognitif tel que le niveau phonologique, puisqu'aucune différence n'est observée en comparaison avec les sujets contrôles, mais qu'elles s'exercent au contraire dans le cadre de hauts niveaux de traitement faisant intervenir des informations lexicales, sémantiques et syntaxiques. On peut donc postuler l'utilisation par les patients implantés de stratégies essentiellement prédictives et intégratives lors de la lecture labiale, ce qui tend à rejoindre les conclusions de notre étude précédente *« Lecture labiale chez le sujet implanté »*, sect. 12.3, p. 268.

Intégration de la parole audiovisuelle

Stratégies audiovisuelles employées par les patients implantés cochléaires

Nous avons démontré dans notre étude précédente (*« Intégration audiovisuelle chez le sujet implanté »*, sect. 12.2, p. 262) que deux mécanismes perceptifs distincts interviennent lors de l'intégration audiovisuelle de la parole chez le sujet implanté cochléaire. D'une part, ces patients développent des capacités de lecture labiale supranormales, probablement par l'intermédiaire de stratégies compensatrices liées à des hauts niveaux de traitement cognitif (stratégies intégratives et prédictrices) ; d'autre part ces patients développent, en plus de ces capacités visuelles purement unimodales, un couplage optimal des informations visuelles avec les informations auditives dégradées fournies par l'implant cochléaire. Toutefois la nature exacte de ce couplage audiovisuel n'est pas connue, et l'on peut se demander dans quelle mesure elle se rapproche du couplage audiovisuel effectué par les sujets normo-entendants.

En utilisant des stimuli de parole audiovisuelle incongruente, dont l'avantage essentiel est de permettre de discriminer les contributions respectives des informations auditives et visuelles lors de l'intégration audiovisuelle, nous avons ainsi pu démontrer une différence entre les mécanismes d'intégration audiovisuelle utilisés par les sujets normo-entendants et les patients implantés cochléaires. Les réponses données dans le cas des stimuli McGurk dentaux montrent en effet que les patients implantés réalisent dans l'immense majorité des cas une intégration à dominante visuelle (percepts /apa/ /aba/ /ama/), tandis que les sujets contrôles réalisent dans une grande majorité des cas une intégration par combinaison des lieux d'articulation auditif et visuel (percepts /apta/ /abda/ /amna/). Les réponses données dans le cas des stimuli McGurk bilabiaux montrent que patients implantés et sujets contrôles adoptent pour la majorité une stratégie de fusion audiovisuelle (percepts /ata/ /ada/ /ana/).

Ainsi, chez le patient implanté cochléaire, la présentation d'un /ama/ visuel avec un /ata/ auditif aboutit la plupart du temps au percept /apa/, contrairement aux sujets normo-entendants qui perçoivent la plupart du temps un /apta/ (voir à ce sujet *« Intégration perceptive audio-visuo-motrice de la parole »*, § 3.6.3, p. 69). On peut interpréter cette différence en faisant l'hypothèse qu'en cas d'audition normale, la présence du /t/ dans le percept audiovisuel est due à la forte pertinence des informations auditives, attestant de façon quasi-certaine de la prononciation d'un phonème postérieur /t/. Au contraire, chez les patients implantés, cette

pertinence est considérablement réduite par la dégradation du signal auditif, et la probabilité de présence du phonème postérieur /t/ (auditif) devient faible devant la probabilité de présence du phonème antérieur /p/ (visuel), ce qui provoque la disparition du /t/ dans le percept audiovisuel final.

Par ailleurs, quel que soit le type de stimulus incongruent utilisé (dental ou bilabial), l'analyse des réponses non audiovisuelles (c'est-à-dire à dominante soit visuelle soit auditive) montre que les sujets normo-entendants ont systématiquement plus tendance à privilégier les réponses à dominante auditive que les patients implantés, tandis que les patients implantés accordent systématiquement une plus grande importance aux réponses à dominante visuelle.

Par conséquent, l'utilisation de stimuli de parole audiovisuelle incongruente permet de mettre en évidence une évaluation différente de la pertinence des informations unimodales entre les sujets normo-entendants et les patients implantés. Au cours de l'intégration audiovisuelle, les patients implantés accorderaient donc un poids très important aux informations visuelles pour obtenir des informations sur le lieu d'articulation, en déséquilibrant l'intégration audiovisuelle en faveur de la vision. Au contraire, les sujets normo-entendants accorderaient des poids équilibrés aux informations auditives et visuelles, avec toutefois un léger avantage en faveur de l'audition (en effet l'effet McGurk n'est pas obtenu dans 100% des cas chez les sujets normo-entendants).

Intégration audiovisuelle phonétique chez les patients implantés cochléaires

L'analyse détaillée des réponses effectuées par les patients implantés pour les stimuli audiovisuels incongruents montre que chez les patients implantés : 1° la perception du voisement et de la nasalité est gouvernée par l'audition restituée par l'implant cochléaire, la vision ne fournissant quasiment aucune information pour ces indices articulatoires ; 2° la perception du lieu d'articulation est fortement orientée par la vision, au détriment des informations fournies par l'implant cochléaire.

Si l'on suppose que ce sont les mêmes mécanismes perceptifs qui gouvernent l'intégration audiovisuelle des stimuli congruents ou incongruents, alors on peut faire la prédiction suivante : la répartition statistique des réponses données pour un stimulus audiovisuel incongruent particulier devrait être identique à la répartition des réponses correspondant au stimulus audiovisuel congruent possédant : 1° le même voisement et la même nasalité que la composante auditive du stimulus incongruent ; 2° le même lieu d'articulation que la composante visuelle du stimulus incongruent.

Pour approfondir cette hypothèse, nous avons effectué une analyse de corrélation entre les réponses données par les patients implantés pour chaque stimulus audiovisuel congruent et chaque stimulus audiovisuel incongruent. Pour chaque couple de stimuli, nous avons calculé l'écart moyen (RMS) entre les deux distributions de réponses. Le résultat de cette analyse (voir fig. 12.10, p. 281) montre une parfaite adéquation avec notre prédiction : pour chaque stimulus audiovisuel incongruent, le stimulus audiovisuel congruent qui aboutit à la répartition de réponses la plus proche (qui correspond donc au plus petit écart-type) est celui qui possède le voisement et la nasalité de la partie auditive du stimulus incongruent, ainsi que le lieu d'articulation de la partie visuelle du stimulus incongruent.

Cette observation peut être retrouvée sur les matrices de confusion (voir fig. 12.8, p. 276) ; par exemple la colonne de réponses correspondant au stimulus incongruent /aga/ visuel + /apa/ auditif possède une répartition particulièrement semblable à la colonne de réponses correspondant au stimulus congruent /ata/. On peut en déduire que « /ata/ audiovisuel » est ce qui se rapproche le plus perceptivement de « /aga/ visuel + /apa/ auditif » pour les patients implantés cochléaires. Il est essentiel de noter que ces observations sont tout à fait cohérentes avec l'hypothèse d'une représentation interne de la parole perçue sous forme de commandes articulatoires correspondant

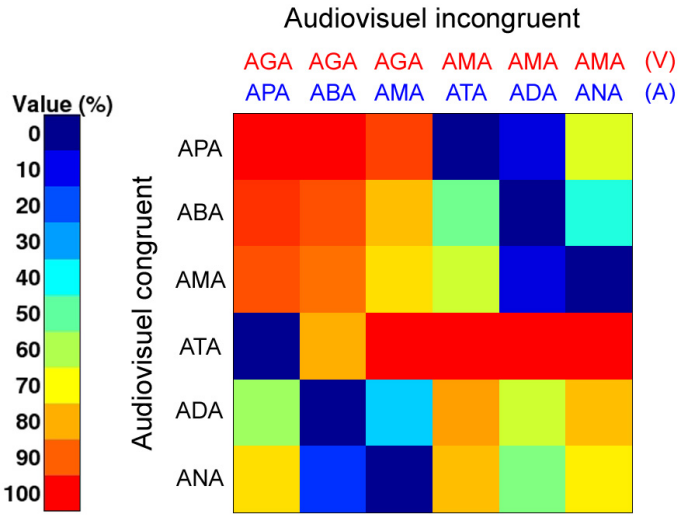


FIG. 12.10 – Comparaison des profils de réponses des patients implantés cochléaires (n = 33) pour les stimuli audiovisuels congruents et incongruents. Pour chaque stimulus incongruent sont représentés les écarts-type (valeur RMS) entre le profil de réponse pour le stimulus incongruent et les profils de réponse pour les différents stimuli congruents. Dans chaque colonne les écarts-type ont été normalisés entre 0 et 100% pour permettre une meilleure lisibilité ; les meilleures corrélations sont donc représentées par les cases de couleur bleu foncé (écart-type normalisé nul).

aux phonèmes compatibles à la fois avec les indices de parole visuels et les indices de parole auditifs (voir *« Intégration perceptive audio-visuo-motrice de la parole »*, § 3.6.3, p. 69).

Evolution longitudinale des mécanismes d'intégration audiovisuelle

Le groupe de 33 patients implantés que nous avons testé est assez hétérogène en ce qui concerne les durées d'utilisation de l'implant. Ces dernières sont en effet réparties entre 0 et 96 mois, avec une moyenne de 21.4 ± 25.8 mois et une médiane de 9 mois. Etant donné le faible effectif et l'importante hétérogénéité du groupe, il est donc difficilement envisageable d'effectuer une subdivision fine en sous-groupes répartis par durée d'implantation en vue d'évaluer l'évolution des capacités d'intégration audiovisuelle phonétique en fonction de la durée de stimulation par l'implant cochléaire.

Toutefois une analyse préliminaire de nos données, bien qu'insuffisante pour être concluante, tend à indiquer que plus la durée d'utilisation de l'implant est importante, mieux les modes d'articulation sont perçus et sont intégrés correctement dans les percepts audiovisuels congruents et incongruents. Par exemple, la présentation d'un /aga/ visuel avec un /apa/ auditif (/p/ n'est ni voisé ni nasalisé) est reconnue dès les premiers mois après implantation comme un /ata/ (/t/ n'est ni voisé ni nasalisé). Au contraire, la présentation d'un /aga/ visuel avec un /aba/ auditif (/b/ est voisé non nasalisé) est d'abord confondue avec /ata/ puis de mieux en mieux identifiée comme /ada/ (/d/ est voisé non nasalisé) à mesure de l'utilisation de l'implant. De façon similaire, la présentation d'un /aga/ visuel avec un /ama/ auditif (/m/ est voisé et nasalisé) est d'abord confondue avec /ata/, puis confondue avec /ata/ et /ada/, pour être finalement identifiée à long terme comme un /ana/ (/n/ est voisé et nasalisé).

Ces constatations préliminaires suggèrent une différence entre les mécanismes d'intégration audiovisuelle existant à court, moyen et long terme après l'implantation cochléaire. Il serait par conséquent intéressant d'en rechercher les corrélats au niveau neuronal, qui pourraient se

traduire par l'observation de différences dans les réseaux d'intégration audiovisuelle de la parole en fonction de la durée d'utilisation de l'implant.

Réseau neuronal d'intégration audiovisuelle

Nos données expérimentales suggèrent une plus grande importance accordée aux signaux visuels lors de l'intégration audiovisuelle de la parole chez les patients sourds implantés cochléaires. Or il est possible d'observer chez ces derniers des activations du cortex visuel associatif en l'absence de toute stimulation visuelle, en réponse aux seuls sons de parole (Giraud *et al.*, 2001; Zatorre, 2001). Ces activations intermodales, inexistantes avant l'implantation cochléaire, se développent progressivement au cours des années qui suivent l'implantation (voir *« Interactions multisensorielles »*, § 9.4.4, p. 203). Par ailleurs, de nombreuses études montrent que les régions activées par la parole visuelle se superposent en partie à celles activées par la parole auditive, en particulier au niveau des aires auditives secondaires (Sams *et al.*, 1991; Calvert *et al.*, 1997; Callan *et al.*, 2001; Bernstein *et al.*, 2002) et primaires (MacSweeney *et al.*, 2000; Molholm et Foxe, 2005; Pekkola *et al.*, 2005; Ruytjens *et al.*, 2006) (voir *« Traitement phonologique visuel de la parole »*, § 4.3.3, p. 84). Cet ensemble de résultats expérimentaux est donc cohérent avec l'hypothèse d'un couplage audiovisuel spécifique se mettant en place au fur et à mesure de l'utilisation de l'implant.

D'un autre côté, les aspects temporels de l'intégration de signaux de parole audiovisuelle incongruents de type McGurk ont été étudiés dans plusieurs études neurofonctionnelles incluant des paradigmes en potentiels évoqués (Sams *et al.*, 1998; Colin *et al.*, 2002, 2004; Musacchia *et al.*, 2006), en mismatch-negativity (Sams *et al.*, 1991; Saint-Amour *et al.*, 2007), en imagerie par résonance magnétique fonctionnelle (Jones et Callan, 2003) et en magnétoencéphalographie (Möttönen *et al.*, 2002) (voir également *« Intégration des informations auditives et visuelles »*, § 4.4.2, p. 85). Ces différentes études concordent pour suggérer une intervention précoce des interactions corticales entre les entrées sensorielles auditive et visuelle, permettant une influence visuelle sur les informations auditives dès le niveau de traitement phonétique (Sams *et al.*, 1998).

En ce qui concerne la localisation précise des réseaux audiovisuels impliqués lors de l'effet McGurk, des recherches récentes suggèrent que la combinaison des informations auditives et visuelles de parole est sous-tendue par un ensemble d'aires localisées au niveau cortical (Macaluso *et al.*, 2004; Skipper *et al.*, 2005; Saito *et al.*, 2005), dont en particulier le cortex pariétal (Jones et Callan, 2003; Hamilton *et al.*, 2006), et même au niveau sous-cortical (Musacchia *et al.*, 2006; Champoux *et al.*, 2006, 2007). L'étude approfondie du comportement et de l'évolution de ces régions après l'implantation cochléaire est donc une clé essentielle pour une compréhension des mécanismes spécifiques d'intégration audiovisuelle utilisés par les patients implantés.

Perspectives concrètes

Nos résultats montrent qu'au niveau phonologique, les patients implantés cochléaires sont tout particulièrement dépendants des indices visuels de la parole concernant la perception du lieu d'articulation et du caractère occlusif/constrictif des consonnes. Malgré ce déséquilibre entre les pertinences accordées par les patients implantés aux indices auditifs et visuels, le mécanisme d'intégration audiovisuelle de la parole est similaire à celui constaté chez les sujets normo-entendants, dans la mesure où il aboutit à la sélection du percept le mieux compatible du point de vue articulatoire avec les indices auditifs et visuels perçus.

Ce déséquilibre audiovisuel en faveur de la vision est à mettre en parallèle avec le haut niveau de lecture labiale conservé par les patients après l'implantation cochléaire (voir *« Perception de la parole chez le sujet implanté »*, sect. 12.1, p. 258) ainsi qu'avec l'optimalité de

l'intégration audiovisuelle de la parole observée chez ces patients (voir *« Intégration audiovisuelle chez le sujet implanté »*, sect. 12.2, p. 262). Ces deux mécanismes jouent probablement un rôle facilitateur essentiel dans les environnements bruyants, pour lesquels les patients implantés rencontrent d'importantes difficultés de compréhension (voir *« Rôle de l'enveloppe temporelle et de la structure fine »*, § 9.2.2, p. 194).

Cet ensemble d'observations contribue donc à souligner l'importance de la bimodalité des signaux de parole durant la période suivant l'implantation cochléaire. En appui avec les résultats de notre étude sur l'entraînement perceptif de sujets normo-entendants avec une simulation acoustique d'implant cochléaire (voir *« Entraînement perceptif avec une simulation d'implant »*, sect. 11.4, p. 249), nous insistons à nouveau sur l'utilisation d'une réhabilitation orthophonique mettant particulièrement l'accent sur une stimulation audiovisuelle, qui fournit des conditions perceptives optimales pour la compréhension de la parole. Nous avons également montré l'intérêt de l'effet McGurk en tant qu'outil d'investigation des mécanismes d'intégration audiovisuelle de la parole, et il est tout à fait envisageable d'utiliser un tel outil comme instrument de diagnostic du degré de récupération des capacités multisensorielles chez le patient implanté cochléaire.

Pour terminer, ces données soulèvent de nombreuses interrogations sur les mécanismes neuronaux sous-tendant l'intégration audiovisuelle de la parole chez les patients implantés cochléaires, en suggérant notamment l'existence de mécanismes différents à court et à long terme après la mise en marche de l'implant. Nous apporterons certains éléments de réponse à ces interrogations dans notre étude en neuroimagerie fonctionnelle (voir *« Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral »*, ch. 13, p. 289).

Bibliographie

Aloufy, S., Lapidot, M. et Myslobodskym, M. (1996). Differences in susceptibility to the "blending illusion" among Native Hebrew and English speakers. *Brain Lang.*, 53(1):51–57.

Bergeson, T. R., Pisoni, D. B. et Davis, R. A. (2003). A longitudinal study of audiovisual speech perception by hearing-impaired children with cochlear implants. *Volta Rev.*, 103:347–370.

Bergeson, T. R., Pisoni, D. B. et Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear Hear.*, 26(2):149–164.

Bernstein, L. E., Auer, Jr., E. T., Moore, J. K., Ponton, C. W., Don, M. et Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, 13:311–315.

Callan, D. E., Callan, A. M., Kroos, C. et Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independant component analysis : A single-sweep EEG case study. *Brain Res. Cogn. Brain Res.*, 10:349–353.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D. et David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276:593–596.

Carpenter, J. et Bithell, J. (2000). Bootstrap confidence intervals : when, which, what ? A practical guide for medical statisticians. *Stat. Med.*, 19(9):1141–1164.

Champoux, F., Paiement, P., Mercier, C., Lepore, F., Lassonde, M. et Gagne, J. P. (2007). Auditory processing in a patient with a unilateral lesion of the inferior colliculus. *Eur. J. Neurosci.*, 25(1):291–297.

Champoux, F., Tremblay, C., Mercier, C., Lassonde, M., Lepore, F., Gagne, J. P. et Theoret, H. (2006). A role for the inferior colliculus in multisensory speech integration. *Neuroreport*, 17(15):1607–1610.

Colin, C., Radeau, M., Soquet, A. et Deltenre, P. (2004). Generalization of the generation of an MMN by illusory McGurk percepts : Voiceless consonants. *Clin. Neurophysiol.*, 115:1989–2000.

Colin, C., Radeau, M., Soquet, A., Demolin, D., Colon, F. et Deltenre, P. (2002). Mismatch negativity evoked by the McGurk-MacDonald effect : A phonetic representation within short-term memory. *Clin. Neurophysiol.*, 113:495–506.

Geers, A., Brenner, C. et Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear Hear.*, 24(1 Suppl):24S–35S.

Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E. et Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, 30(3):657–663.

Goh, W. D., Pisoni, D. B., Kirk, K. I. et Remez, R. E. (2001). Audio-visual perception of sinewave speech in an adult cochlear implant user : a case study. *Ear Hear.*, 22(5):412–419.

Hamilton, R. H., Shenton, J. T. et Coslett, H. B. (2006). An acquired deficit of audiovisual speech processing. *Brain Lang.*, 98(1):66–73.

Hamzavi, J., Baumgartner, W. D., Pok, S. M., Franz, P. et Gstoettner, W. (2003). Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Otolaryngol.*, 123(4):493–498.

Irwin, J. R., Whalen, D. H. et Fowler, C. A. (2006). A sex difference in visual influence on heard speech. *Percept. Psychophys.*, 68(4):582–592.

Johnson, F. M. et Leslie, H. H. (1988). Sex differences in lipreading. *Bull. Psychon. Soc.*, 26(2):106–108.

Jones, J. A. et Callan, D. E. (2003). Brain activity during audiovisual speech perception : an fMRI study of the McGurk effect. *Neuroreport*, 14(8):1129–1133.

Kaiser, A. R., Kirk, K. I., Lachs, L. et Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J. Speech Lang. Hear. Res.*, 46(2):390–404.

Macaluso, E., George, N., Dolan, R., Spence, C. et Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech : a PET study. *Neuroimage*, 21(2):725–732.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., Williams, S. C., Woll, B. et Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise : An event-related fMRI study. *Neuroreport*, 11:1729–1733.

Molholm, S. et Foxe, J. J. (2005). Look "hear", primary auditory cortex is active during lip-reading. *Neuroreport*, 16:123–124.

Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer, Jr., E. T., Fisher, L. et Bernstein, L. E. (2005). Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol. Neurotol.*, 26(4): 649–654.

Möttönen, R., Krause, C. M., Tippana, K. et Sams, M. (2002). Processing of visual speech in the human auditory cortex. *Brain Res. Cogn. Brain Res.*, 13:417–425.

Musacchia, G., Sams, M., Nicol, T. et Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Exp. Brain Res.*, 168:1–10.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A. et Sams, M. (2005). Primary auditory cortex activation by visual speech : An fMRI study at 3T. *Neuroreport*, 16:125–128.

Pelizzone, M., Cosendai, G. et Tinembart, J. (1999). Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for Ineraid implanted subjects. *Ear Hear.*, 20(3):228–237.

R (2007). The r project for statistical computing. Site internet. http://www.r-project.com.

Ruytjens, L., Albers, F., van Dijk, P., Wit, H. et Willemsen, A. (2006). Neural responses to silent lipreading in normal hearing male and female subjects. *Eur. J. Neurosci.*, 24(6):1835–1844.

Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W. et Foxe, J. J. (2007). Seeing voices : High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia*, 45(3):587–597.

Saito, D. N., Yoshimura, K., Kochiyama, T., Okada, T., Honda, M. et Sadato, N. (2005). Cross-modal binding and activated attentional networks during audio-visual speech integration : a functional MRI study. *Cereb. Cortex*, 15(11):1750–1760.

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T. et Simola, J. (1991). Seeing speech : visual information from lip movements modifies activity in the human auditory cortex. *Neurosci. Lett.*, 127:141–145.

Sams, M., Manninen, P., Surakka, V., Helin, P. et Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences : Effects of word meaning and sentence context. *Speech Comm.*, 26:75–87.

Skipper, J. I., Nusbaum, H. C. et Small, S. L. (2005). Listening to talking faces : motor cortical activation during speech perception. *Neuroimage*, 25(1):76–89.

Strelnikov, K. (2007). Activation-verification in continuous speech processing. Interaction of cognitive strategies as a possible theoretical approach. *J. Neurolinguist.*

Tyler, R. S., Gantz, B. J., Parkinson, A. J., Woodworth, G. G., Lowder, M. W. et Parkinson, W. S. (1997). Initial comparison between the Clarion, Nucleus and Ineraid cochlear implants. *Adv. Otorhinolaryngol.*, 52:274–278.

Watson, C. S., Qiu, W. W., Chamberlain, M. M. et Li, X. (1996). Auditory and visual speech perception : confirmation of a modality-independent source of individual differences in speech recognition. *J. Acoust. Soc. Am.*, 100(2 Pt 1):1153–1162.

Zatorre, R. J. (2001). Do you see what I'm saying ? interactions between auditory and visual cortices in cochlear implant users. *Neuron*, 31(1):13–14.

Chapitre 13

Traitement cortical de la parole audiovisuelle chez l'implanté cochléaire unilatéral

Sommaire

13.1 Matériel et méthodes 291
13.2 Activations cérébrales chez les patients implantés 294
13.2.1 Différence entre implantation droite et gauche 294
13.2.2 Activations visuelles 296
13.2.3 Activations audiovisuelles 302
13.2.4 Evolution longitudinale des activations 305
13.2.5 Récapitulatif des activations cérébrales 307
13.3 Activations cérébrales chez les sujets contrôles 308
13.3.1 Activations visuelles 308
13.3.2 Activations audiovisuelles 310
13.3.3 Récapitulatif des activations cérébrales 311
13.4 Différences d'activité patients/contrôles 312
13.4.1 Différences d'activité en stimulation visuelle 312
13.4.2 Différences d'activité en stimulation audiovisuelle 313
13.4.3 Récapitulatif des différences d'activations cérébrales 318
13.5 Etude de corrélation des activations cérébrales 319
13.5.1 Méthodes 319
13.5.2 Corrélation entre données comportementales et activations cérébrales . . 319
13.5.3 Connectivité cérébrale 319
13.6 Discussion 320
13.6.1 Réorganisation cérébrale du traitement visuel de la parole 320
13.6.2 Réorganisation cérébrale du traitement audiovisuel de la parole 322
13.6.3 Rôle de l'aire temporale de la voix chez le patient implanté 323
13.6.4 Rôle du réseau audio-visuo-moteur chez le patient implanté 324
13.6.5 Evolution post-implantation des stratégies de traitement de la parole . . 326
Bibliographie 329

Introduction

Nos données psychophysiques, appuyées par les données de la littérature scientifique, mettent en évidence une modification du traitement audiovisuel de la parole en fonction de la durée d'utilisation de l'implant cochléaire, en indiquant une stabilisation de ce traitement autour de sept mois après l'implantation cochléaire, lorsque les performances auditives ont dépassé le seuil de 80% et commencent à se stabiliser. Par ailleurs les observations comportementales indiquent le rôle primordial joué par les indices visuels de parole dans cette réorganisation des mécanismes audiovisuels.

Toutefois les mécanismes de réorganisation corticale prenant place après l'implantation cochléaire sont peu connus à l'heure actuelle, et très peu d'études se sont intéressées à la plasticité des réseaux neuronaux d'intégration audiovisuelle de la parole chez le patient sourd implanté cochléaire. Afin d'approfondir cette piste de recherche nous avons donc mis au point un protocole de suivi longitudinal des patients implantés en imagerie fonctionnelle par tomographie par émission de positrons.

13.1 Matériel et méthodes

Sujets

Nous avons réalisé une analyse prospective longitudinale non randomisée sur 10 patients adultes présentant une surdité profonde bilatérale postlinguale et ayant subi une implantation cochléaire unilatérale. Le groupe de sujets est composé de 7 femmes et 3 hommes entre 35 et 81 ans, avec une moyenne d'âge de 53.9 ans. Les étiologies sont très variables dans le groupe de sujets, principalement infectieuses et génétiques (méningite, otite chronique, otosclérose, neurinome acoustique, otospongiose, etc.) La majorité des patients ont reçu un implant cochléaire Nucleus CI-24 avec l'une des stratégies de codage RCA ou RCS. Le groupe de sujets contrôles est quant à lui composé de 6 hommes entre 20 et 49 ans.

Cette étude a reçu un avis favorable des comités éthiques locaux (le 13 décembre 2004 à Toulouse par délibération du Comité consultatif de protection des personnes se prêtant à la recherche biomédicale). Tous les patients ont donné leur accord écrit après complète information sur le protocole expérimental de cette étude.

Patients implantés						
Sujet	**Age**	**Sexe**	**Côté**	**T0**	**T1**	**Implant**
IC02	81	F	D	0m	X	Nucleus CI24RCS
IC03	39	M	D	0m	10m	Nucleus CI24RCS
IC04	39	F	D	0m	4m	Nucleus CI24RCS
IC06	57	F	G	1m	3m	Nucleus CI24RCA
IC07	69	M	G	0m	10m	MEDEL
IC08	39	M	G	0m	X	Nucleus CI24RCA
IC09	62	F	G	0m	4m	Clarion
IC10	64	F	G	0m	9m	Advanced Bionics
IC11	54	F	D	0m	6m	Nucleus CI24
IC12	35	F	D	0m	6m	Nucleus CI24

Sujets contrôles		
Sujet	**Age**	**Sexe**
NH01	49	M
NH02	46	M
NH03	26	M
NH04	38	M
NH05	26	M
NH06	20	M

TAB. 13.1 – Patients implantés et sujets contrôles

Stimuli

Nos données comportementales démontrant l'intérêt particulier représenté par le rôle de la vision lors de l'intégration audiovisuelle de la parole chez les patients implantés cochléaires, nous avons défini des conditions de stimulation par de la parole dans les modalités visuelle et audiovisuelle. En plus de ces stimulations par de la parole, nous avons utilisé deux conditions contrôles avec des stimulations non linguistiques, afin d'effectuer de façon appropriée l'analyse des profils d'activations cérébrales. En raison du caractère longitudinal de notre étude et des limitations médicales des doses radioactives admises, nous avons dû nous restreindre aux stimulations par de la parole visuelle et audiovisuelle, malgré l'important intérêt potentiel d'une comparaison avec une stimulation uniquement auditive.

CONDITIONS VISUELLE ET AUDIOVISUELLE (V ET AV) : Les stimuli sont tous des mots dissyllabiques faisant partie des listes normalisées de Fournier. Ces mots ont été répartis en trois listes de 20 mots chacune (L1-L2-L3). Ces listes ont toutes été équilibrées en fréquence d'utilisation

lexicale, comptabilisée en nombre d'occurences par million de mots d'après la base de données Brulex (Content *et al.*, 1990), en structure phonétique (répartition voyelles antérieures/voyelles postérieures) et en structure syllabique (CV/CVC/CCV). L'enregistrement des prononciations de chaque mot s'est déroulé suivant les modalités décrites dans « *Génération des stimuli* », § 10.2.1, p. 218.

Les stimuli sont présentés en condition auditive seule et audiovisuelle. Deux acquisitions sont réalisées par séance pour chaque condition (visuelle et audiovisuelle). Afin de tester l'identification correcte des mots, nous ajoutons à chaque groupe de 20 mots un groupe de 20 "non-mots", réalisés par inversion temporelle des 20 mots originaux, puis nous présentons cet ensemble de 40 stimuli mélangés dans un ordre aléatoire différent pour chaque présentation et pour chaque sujet. Les patients reçoivent l'instruction d'indiquer pour chaque stimuli s'ils reconnaissent ou pas un mot français.

Après rajout des non-mots puis mélange aléatoire, chacun des trois groupes de 40 stimuli est enregistré sous forme visuelle seule (sans le son) et sous forme audiovisuelle, ce qui donne six films de 40 stimuli comprenant chacun 20 mots et 20 non-mots : trois en condition de lecture labiale (V1/V2/V3) et trois en condition audiovisuelle (AV1/AV2/AV3).

En accord avec l'orthophoniste chargée du suivi post-implantation des patients, les listes L1 et L3 (stimuli V1/AV1 et V3/AV3) correspondent à des mots qui ne sont pas travaillés lors de la réhabilitation orthophonique, tandis que la liste L2 (stimuli V2/AV2) correspond à des mots travaillés. Le but de cette différenciation est de pouvoir étudier les conséquences de l'apprentissage au niveau comportemental et neurologique.

Référence motrice (RM) : Afin de détecter les activations particulières aux réseaux corticaux impliqués dans la perception du langage, nous définissons également des stimuli de "référence motrice", destinés à mettre en valeur les activations non-spécifiques du langage qui sont induites par nos stimuli visuels et audiovisuels. Deux acquisitions sont réalisées par séance en condition de référence motrice. Ces stimuli sont définis de la façon suivante : 40 non-mots sont présentés en condition visuelle seule (image inversée temporellement) avec, pour 50% des stimuli choisis de façon aléatoire, un affichage inversé de l'image (inversion du haut et du bas). Les patients reçoivent l'instruction d'indiquer pour chaque stimuli si le visage est affiché à l'envers ou pas, sans se préoccuper de reconnaître ce qui est prononcé. L'objectif recherché est de détecter les activations corticales non spécifiques du langage, en particulier celles liées à la motricité (mouvements de réponse de la main) et à la présentation de l'image du visage de l'orthophoniste dans un contexte non-linguisitique, afin de s'affranchir de ces activations non-spécifiques lors de l'analyse des activations en lecture labiale et en perception audiovisuelle de la parole.

Repos (R) : Deux acquisitions de repos sont également planifiées par séance, afin de disposer d'une référence sur l'activation corticale de repos spécifique à chaque sujet. L'acquisition de repos se fait dans le silence, le patient recevant l'instruction de fermer les yeux pendant la durée du repos. Un masque de relaxation opaque est utilisé comme précaution supplémentaire pour éviter toute stimulation visuelle.

Procédure

L'examen consiste en une première acquisition de calibrage (scan de transmission) d'une durée de 10 min, suivie de 8 acquisitions de 180 secondes séparées par des temps de décroissance

radioactive d'environ 6 minutes. Ces acquisitions se succédent dans l'ordre indiqué dans le tableau 13.2. La durée totale de l'examen, avec inclusion des temps de repos, est au maximum de 2 heures.

n°	Etape	Stimulation auditive	Stimulation visuelle	Durée
0	Calibrage	—	—	10 min
1	Repos (A)	—	—	3(+6) min
2	RM (A)	—	Visages endroit/envers	3(+6) min
3	V (A)	—	Lecture labiale sans/avec inversion	3(+6) min
4	V (B)	—	Lecture labiale sans/avec inversion	3(+6) min
5	AV (A)	Mots sans/avec inversion	Lecture labiale sans/avec inversion	3(+6) min
6	AV (B)	Mots sans/avec inversion	Lecture labiale sans/avec inversion	3(+6) min
7	RM (B)	—	Visages endroit/envers	3(+6) min
8	Repos (B)	—	—	3 min

TAB. 13.2 – Déroulement d'une séance de tomographie

Deux séances de tomographie sont planifiées pour chacun des patients implantés. La tomographie de référence (T0) a lieu dans la semaine suivant la mise en marche de l'implant cochléaire, et comprend les films V1/AV1 et V2/AV2. Les patients sont ensuite suivis régulièrement en consultation orthophonique dans le cadre de leur réhabilitation post-implantation. Cette réhabilitation inclut les mots des séries L1 (V1/AV1) et L2 (V2/AV2) mais exclut les mots de la série L3 (V3/AV3).

Une tomographie de contrôle longitudinal (T1) est planifiée dès que les patients dépassent un score de 80% d'intelligibilité de la parole en audition seule (implant cochléaire sans l'aide de la lecture labiale). Dans le cas où ils ne dépassent pas ce score de 80%, la tomographie de contrôle est réalisée un an après la tomographie de référence. Cette séance de contrôle comprend les films V3/AV3 (non travaillés en consultation orthophonique) et V2/AV2 (travaillés en consultation). Pour cette tomographie de contrôle, seules les données d'activation de 8 patients sont disponibles (4 implantés à droite et 4 implantés à gauche). En effet, les patients IC02 et IC08 ont été exclus de l'étude respectivement pour déviation du protocole expérimental et pour explantation cochléaire.

13.2 Activations cérébrales chez les patients implantés

13.2.1 Différence entre implantation droite et gauche

Notre groupe de 10 patients implantés cochléaires inclut 5 patients implantés à droite et 5 patients implantés à gauche. Les implications du côté d'implantation du point de vue de la réorganisation corticale post-opératoire ne sont pas connues, aucune étude en neuroimagerie fonctionnelle ne s'étant penchée sur la question.

Etant donné les différences fonctionnelles qui distinguent les deux hémisphères cérébraux, en particulier au niveau de l'intégration audiovisuelle de la parole (*« Intégration des informations auditives et visuelles »*, § 4.4.2, p. 85), du traitement du langage (*« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91) et des réseaux de haut traitement cognitif (*« Réseau frontal inférieur droit »*, § 4.5.4, p. 95), il est tout à fait envisageable que la réorganisation corticale faisant suite à l'implantation cochléaire soit en partie conditionnée par le côté choisi pour l'implantation. Par conséquent, toute analyse rigoureuse doit distinguer les implantés à droite des implantés à gauche, afin de détecter d'éventuelles différences.

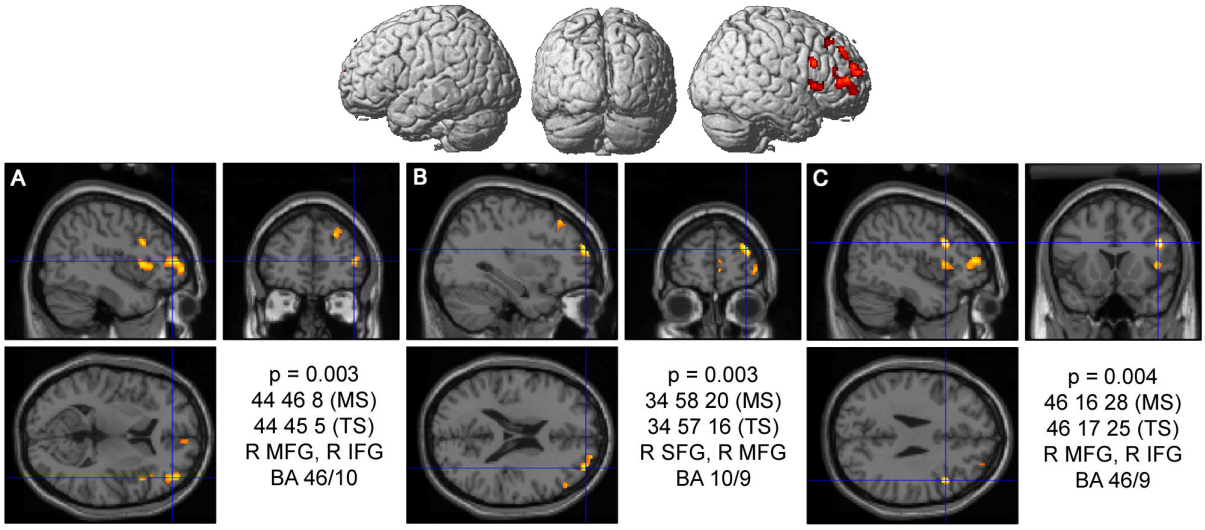


FIG. 13.1 – Différences d'activations cérébrales observées à T0 entre les 5 patients implantés à gauche et les 5 patients implantés à droite (G > D) en condition de parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

Pour chaque condition de stimulation, nous avons donc analysé de façon détaillée les différences d'activation en fonction du côté d'implantation. À T0, la seule activité différentielle statistiquement significative au seuil de 5% est une suractivation par la parole visuelle et audiovisuelle des régions du lobe frontal droit chez les implantés du côté gauche (voir fig. 13.1, p. 294 et fig. 13.2, p. 295). Chez les implantés à gauche testés après récupération perceptive (T1), cette suractivation du lobe frontal droit disparaît, tandis qu'on observe l'apparition d'une suractivation du gyrus fusiforme droit (voir fig. 13.3, p. 295). Chez les implantés à droite testés à T1, on voit apparaître une suractivation de la partie dorsale de l'aire de Broca (gyrus frontal moyen postérieur dorsal ; voir fig. 13.3, p. 295)

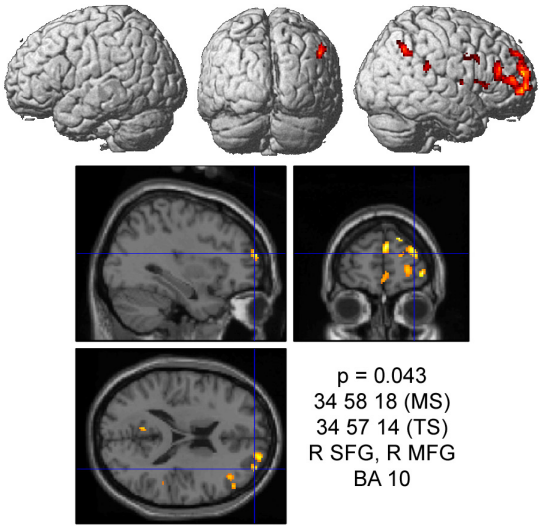


FIG. 13.2 – Différences d'activations cérébrales observées à T0 entre les 5 patients implantés à gauche et les 5 patients implantés à droite (G > D) en condition de parole audiovisuelle (contraste AV-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

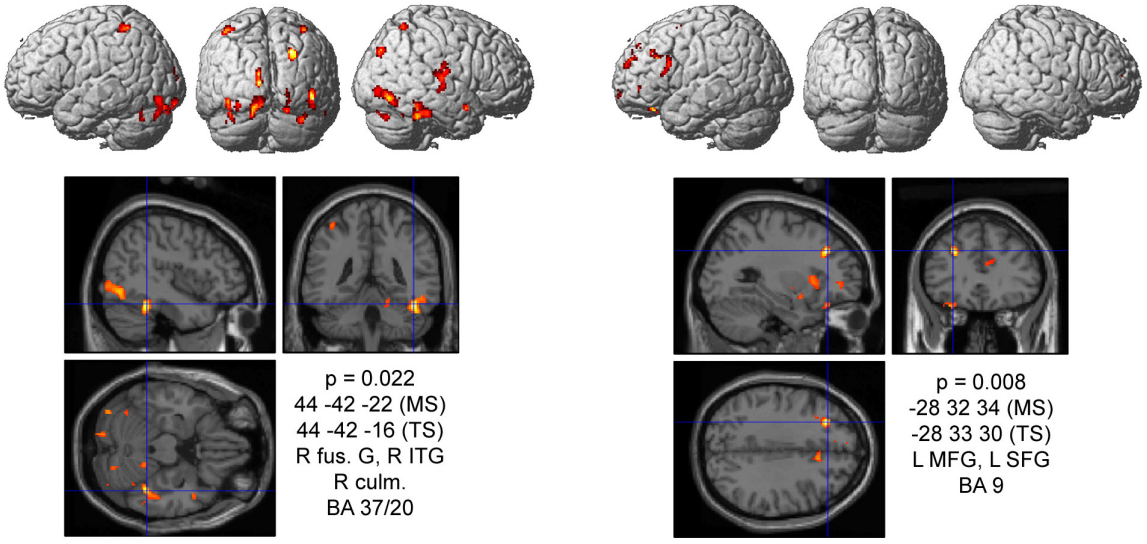


FIG. 13.3 – Différences d'activations cérébrales observées à T1 entre les 4 patients implantés à gauche et les 4 patients implantés à droite (G > D et D > G) en condition de parole visuelle avec activations spécifiques du langage (contraste V-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

13.2.2 Activations visuelles

Contraste RM-R

Les foyers d'activations cérébrales observés sur notre groupe de 10 patients implantés cochléaires peu après la mise en marche de leur implant (T0) lors de la présentation de visages dans une tâche non linguistique sont essentiellement localisés dans les aires visuelles (voir fig. 13.4, p. 296 et fig. 13.5, p. 297). On observe en particulier une forte activation bilatérale dans les aires visuelles associatives, au niveau de la partie postérieure des gyri occipitaux moyen et inférieur ainsi que dans la partie postérieure du gyrus fusiforme, du gyrus lingual et du cuneus (A et B, fig. 13.4 et 13.5).

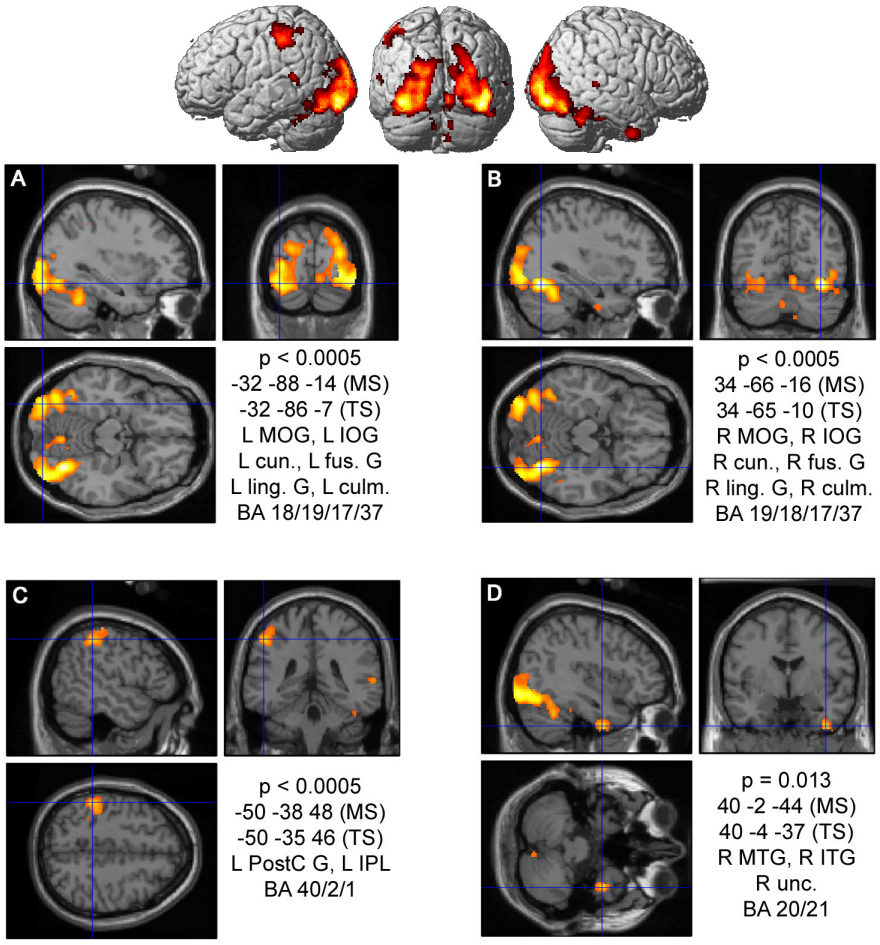


FIG. 13.4 – Activations cérébrales de groupe observées pour les patients implantés testés à T0 (n = 10). Présentation des visages sans contexte linguistique (contraste RM-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

On peut également noter la présence attendue d'un foyer d'activation localisé dans la partie dorso-latérale du gyrus postcentral gauche et du lobule intrapariétal, ce qui correspond aux activations motrices liés aux mouvements des doigts de la main droite lors des réponses comportementales (C, fig. 13.4 et 13.5). Enfin il est également intéressant d'observer une activation latéralisée à droite dans la partie antérieure du gyrus temporal inférieur (D, fig. 13.4). Cette activation sera discutée de façon approfondie dans la suite.

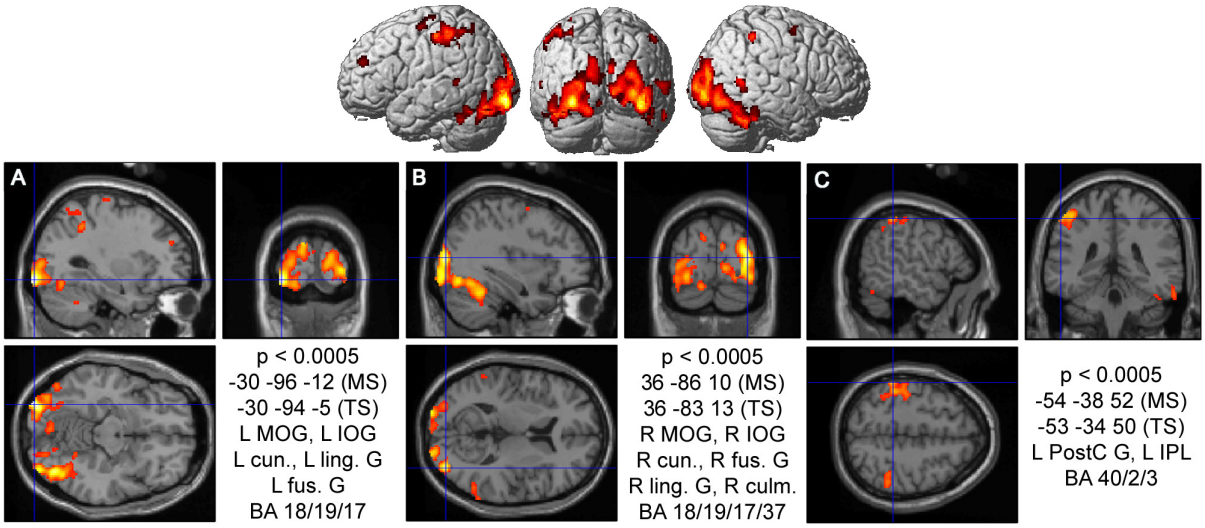


FIG. 13.5 – Activations cérébrales de groupe observées pour les patients implantés testés à T1 (n = 8). Présentation des visages sans contexte linguistique (contraste RM-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

Contraste V-R

Les foyers d'activations cérébrales observés chez les patients implantés cochléaires lors de la présentation de parole visuelle impliquent des régions principalement réparties sur les lobes occipitaux, temporaux et frontaux (voir fig. 13.6, p. 298, fig. 13.7, p. 299 et fig. 13.8, p. 300). Comme dans le cas de la présentation de visages dans une tâche non-linguistique (contraste RM-R), on observe une forte activation bilatérale dans les aires visuelles associatives, au niveau de la partie postérieure des gyri occipitaux moyen et inférieur ainsi que dans la partie postérieure du gyrus fusiforme, du gyrus lingual et du cuneus, aussi bien à T0 qu'à T1 (A–D, fig. 13.6 et A–B, fig. 13.8). On peut également retrouver la présence attendue d'un foyer d'activation localisé dans la partie dorso-latérale du gyrus postcentral gauche et du lobule intrapariétal, ce qui correspond aux activations motrices liés aux mouvements des doigts de la main droite lors des réponses comportementales (G, fig. 13.7). Bien qu'apparente dans le groupe à T1, cette activation passe légèrement en-dessous du seuil de significativité statistique de 5%.

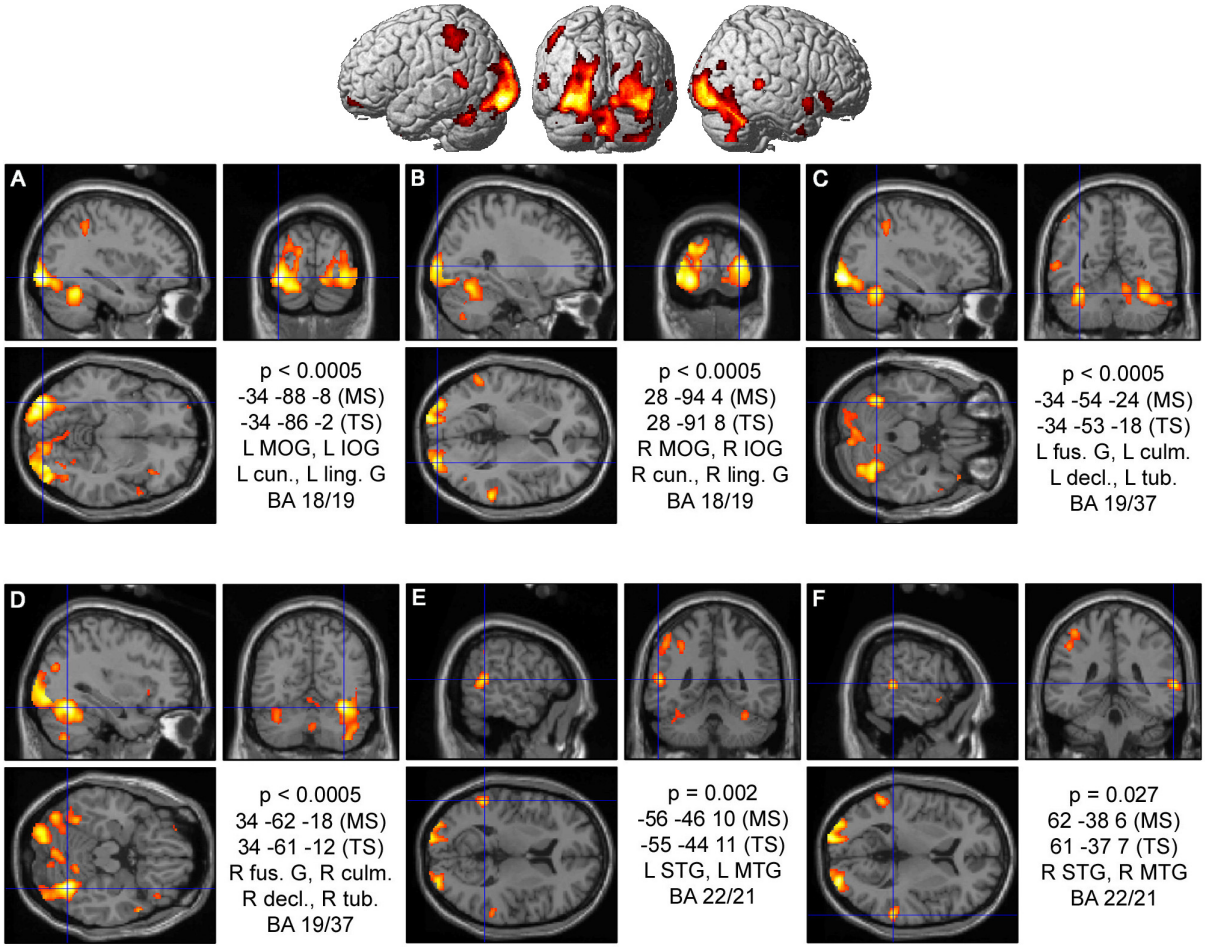


FIG. 13.6 – Partie I. Activations cérébrales de groupe observées pour les patients implantés testés à T0 (n = 10). Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

Parmi les aires qui ne sont pas directement impliquées dans la vision ou la motricité, on trouve à T0 un foyer d'activation latéralisé à droite dans la partie antérieure du gyrus temporal supérieur (I, fig. 13.7), ce foyer pouvant correspondre à l'une des aires cérébrales sélectives à l'identité de la voix. Ce foyer, bien qu'apparent à T1, n'est toutefois plus significativement activé

après récupération perceptive.

De façon particulièrement intéressante, on observe également une activation bilatérale des aires auditives associatives situées dans la partie postérieure du sillon temporal supérieur, proche du gyrus angulaire (E–F, fig. 13.6), révélatrice d'interactions entre la vision et l'audition. Parmi ces deux activations, seul le foyer situé dans le lobe temporal gauche demeure significativement activé à T1 (C, fig. 13.8).

Au niveau du lobe frontal, on observe à T0 un foyer d'activation dans la partie postérieure du gyrus frontal inférieur droit (homologue de l'aire de Broca) (J, fig. 13.7), ainsi qu'un foyer d'activation dans la partie antéro-ventrale du lobe frontal gauche (K, fig. 13.7). On retrouve également des foyers d'activation frontaux bilatéraux à T1, avec cette fois un niveau de significativité plus élevé et des localisations sensiblement différentes : le foyer frontal droit observé à T1 est plus antérieur (F, fig. 13.8) tandis que le foyer frontal gauche à T1 est plus latéro-postérieur, au niveau de la partie ventrale de l'aire de Broca (E, fig. 13.8).

Pour terminer, on remarque à T0 une activation du cervelet au niveau de la partie médiale droite du culmen (H, fig. 13.7). Un foyer similaire est observé à T1, avec toutefois une localisation plus postérieure (D, fig. 13.8).

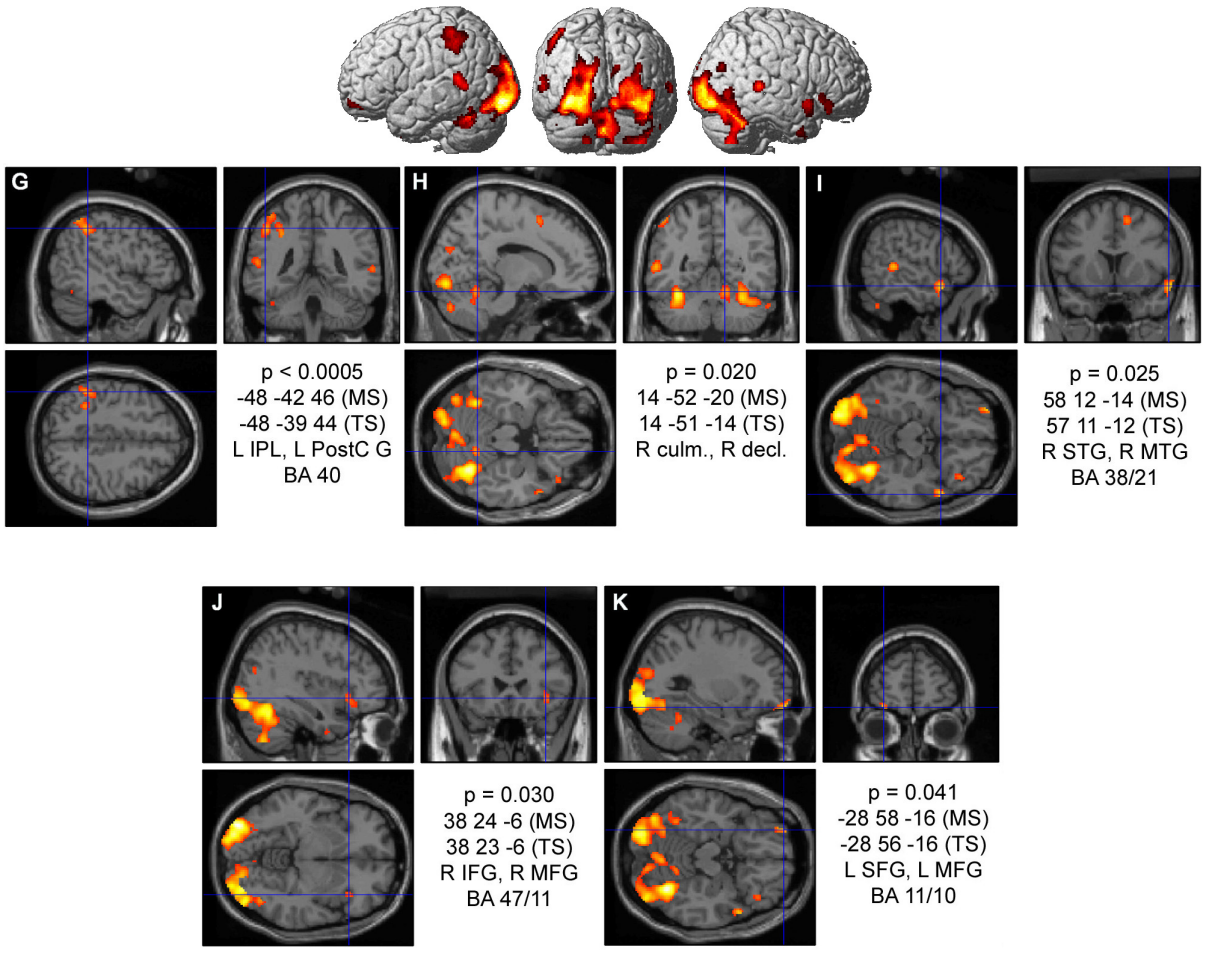


FIG. 13.7 – Partie II. Activations cérébrales de groupe observées pour les patients implantés testés à T0 (n = 10). Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

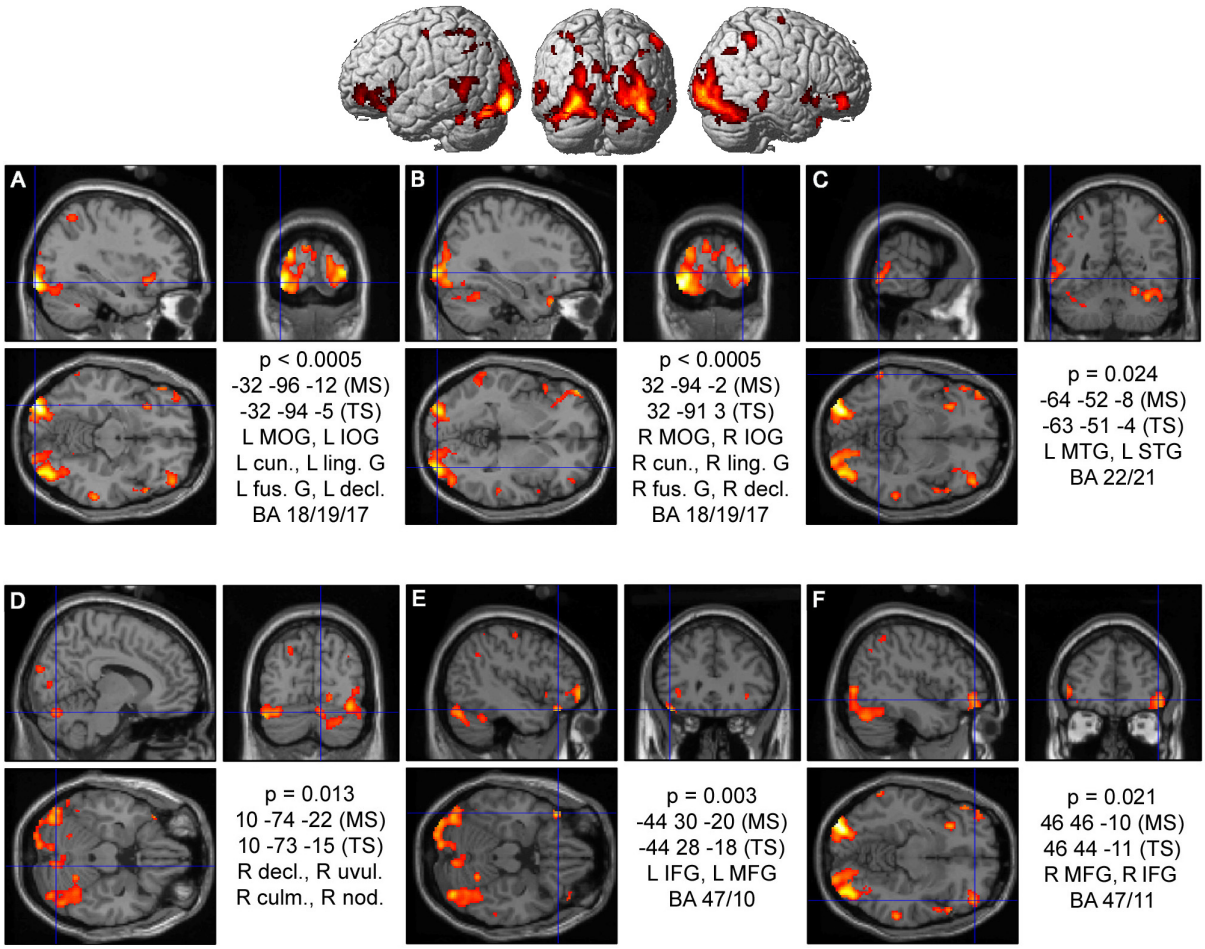


FIG. 13.8 – Activations cérébrales de groupe observées pour les patients implantés testés à T1 (n = 8). Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

Contraste V-RM

Les foyers d'activations cérébrales observés chez les patients implantés cochléaires lors d'une tâche linguistique de lecture labiale contrastée avec une présentation de visages en contexte non-linguistique impliquent des régions localisées dans le cervelet et le lobe frontal (voir fig. 13.9, p. 301 et fig. 13.10, p. 301).

On observe en particulier un foyer d'activation à T0 situé dans la partie médio-occipitale du cervelet, entre la déclive et le pyramis (A, fig. 13.9). On peut également remarquer un foyer d'activation à T0 latéralisé à droite sur la face latérale du cervelet, situé entre le tuber, le pyramis, l'amygdale cérebelleuse et le lobule semi-lunaire inférieur (C, fig. 13.9). Ces deux foyers disparaissent à T1, après récupération perceptive.

Au niveau du lobe frontal, on retrouve à T0 le foyer frontal inférieur postérieur droit (B, fig. 13.9) qui avait déjà pu être observé dans la tâche de lecture labiale contrastée avec le repos (J, fig. 13.7, p. 299). Ce foyer, correspondant à la partie ventrale de l'aire homologue de Broca (à droite), reste apparent à T1 mais ne dépasse plus le seuil de significativité statistique, indiquant une probable déactivation des aires frontales droites après récupération perceptive.

Le foyer frontal inférieur antéro-ventral gauche, remarqué à T0 dans la tâche de lecture labiale contrastée avec le repos (K, fig. 13.7, p. 299), est encore apparent à T0 dans le contraste

avec la condition visuelle de référence (RM) sans toutefois atteindre le seuil de significativité. À T1, on observe au contraire une forte activation de la partie ventrale de l'aire de Broca (fig. 13.10), indiquant une probable réactivation des aires frontales gauches après récupération perceptive.

Comme dans le cas du contraste V-R, on remarque à T0 aussi bien qu'à T1 un foyer de suractivation dans le gyrus temporal supérieur antérieur droit. Bien que cette suractivation reste statistiquement non-significative aussi bien à T0 qu'à T1, il est intéressant d'observer celle-ci à nouveau dans la région cérébrale impliquée dans l'identification de la voix.

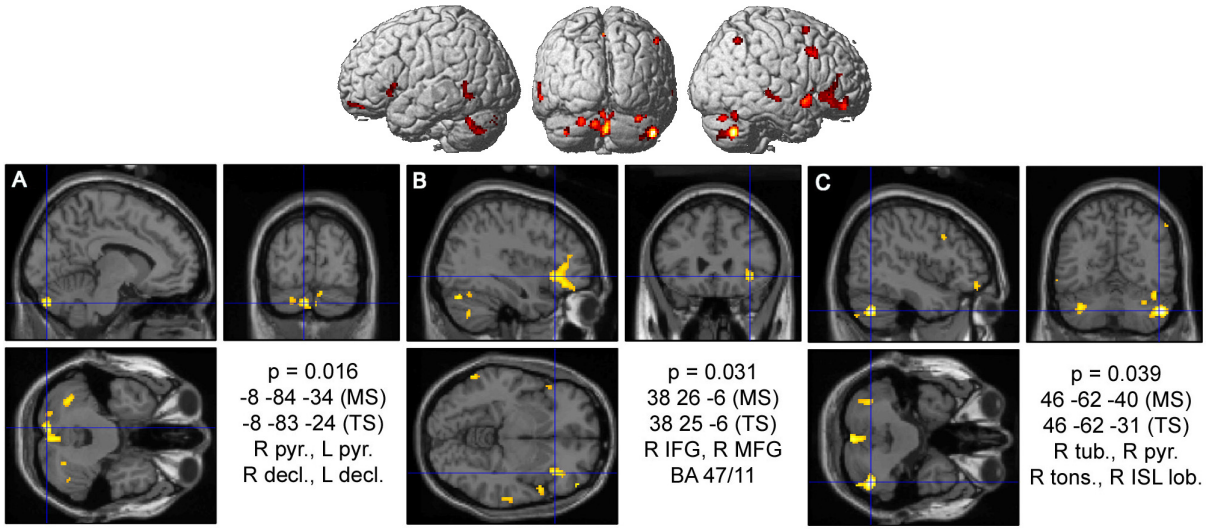


FIG. 13.9 – Activations cérébrales de groupe observées pour les patients implantés testés à T0 (n = 10). Parole visuelle avec activations spécifiques du langage (contraste V-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

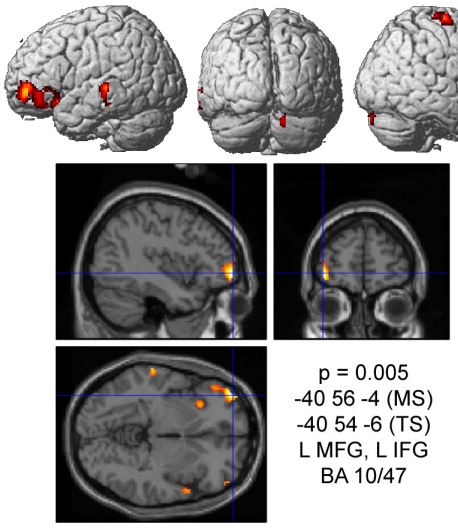


FIG. 13.10 – Activations cérébrales de groupe observées pour les patients implantés testés à T1 (n = 8). Parole visuelle avec activations spécifiques du langage (contraste V-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

13.2.3 Activations audiovisuelles

Contraste AV-RM

Parmi les activations cérébrales observées lors de la perception de parole audiovisuelle (contrastée avec la condition de référence motrice), on peut observer une activation bilatérale des aires auditives situées dans le gyrus temporal supérieur, à T0 aussi bien qu'à T1 (A–B, fig. 13.11, p. 302 et A–C, fig. 13.12, p. 303).

À T0, le foyer auditif droit ($p < 0.0005$) s'étend aux aires primaires et secondaires dans les aires de Brodmann 41/42 et 21/22/38, jusqu'au niveau des régions temporales antérieures droites sélectives à l'identité de la voix (B, fig. 13.11). À T1, on observe une diminution notable du niveau de significativité statistique ($p = 0.001$) accompagnée d'une activation focalisée dans les aires auditives secondaires (BA 21/22/38) au détriment des aires auditives primaires (C, fig. 13.12).

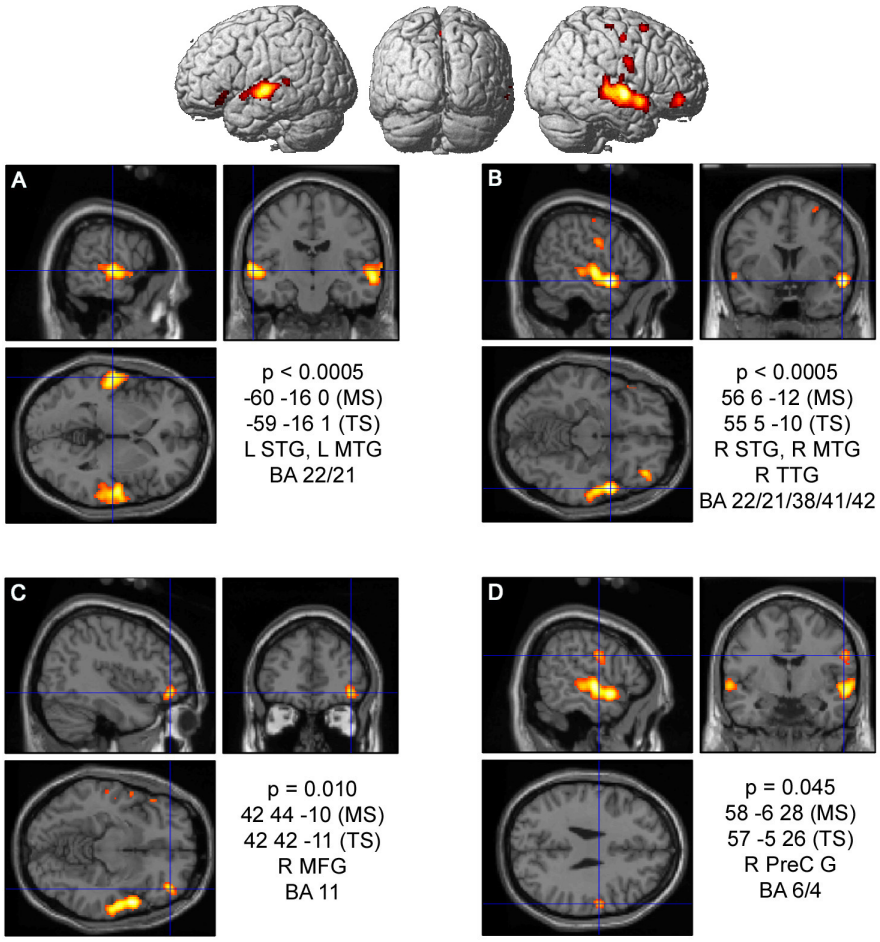


FIG. 13.11 – Activations cérébrales de groupe observées pour les patients implantés testés à T0 ($n = 10$). Parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.001$

Le foyer auditif gauche observé à T0 reste confiné dans les aires auditives associatives BA 21/22 (A, fig. 13.11). Après récupération perceptive (à T1), on peut remarquer une extension des activations auditives gauches en deux foyers (A–B, fig. 13.12), avec une extension vers la partie

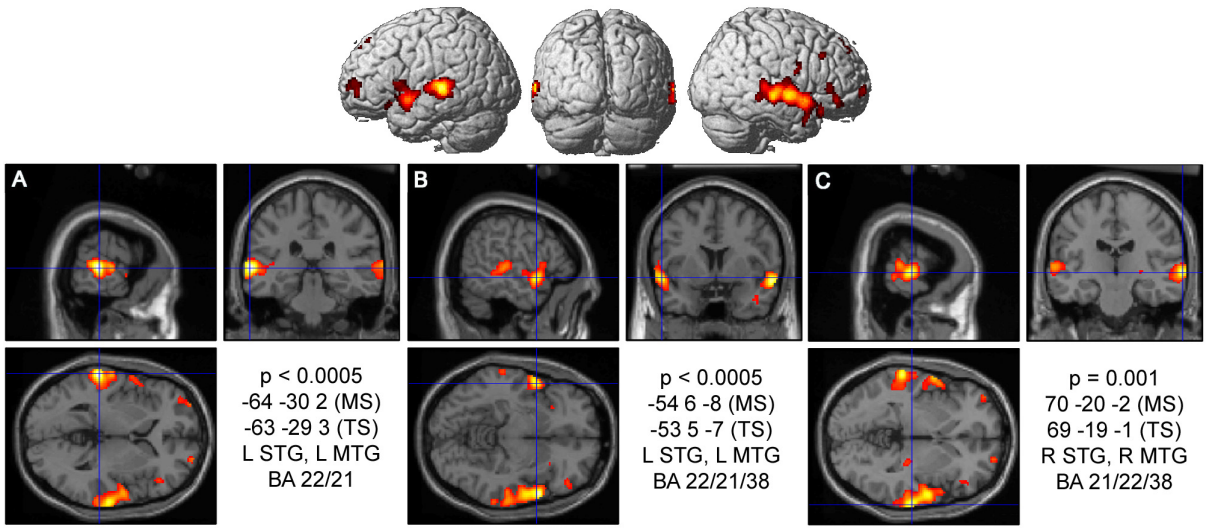


FIG. 13.12 – Activations cérébrales de groupe observées pour les patients implantés testés à T1 (n = 8). Parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.001$

antérieure du gyrus temporal supérieur gauche, jusqu'à l'aire de Brodmann 38.

À T0, on peut également noter deux autres foyers d'activation, le premier situé dans la partie antérieure du sillon frontal inférieur droit, et le second dans la partie moyenne du gyrus précentral droit (C–D, fig. 13.11). Ces foyers disparaissent à T1, après récupération perceptive.

Contraste AV-V

Les suractivations cérébrales observées lors de la perception de parole audiovisuelle (contrastée avec la parole visuelle) sont regroupées bilatéralement dans la partie moyenne du gyrus temporal supérieur (voir fig. 13.13, p. 304 et fig. 13.14, p. 304). Les foyers d'activations associés se trouvent bilatéralement dans les aires de Brodmann 21 et 22 (cortex auditif associatif), avec une extension à droite à l'aire 42, à proximité du cortex auditif primaire, observée à T0 (mais pas à T1).

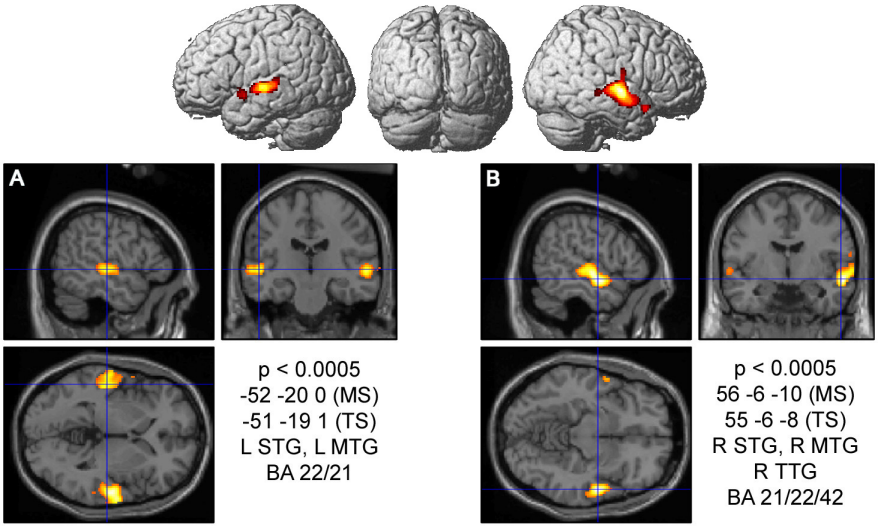


FIG. 13.13 – Activations cérébrales de groupe observées pour les patients implantés testés à T0 (n = 10). Comparaison entre les activations par de la parole audiovisuelle et visuelle (contraste AV-V). Volumes supérieurs à 30 voxels avec $p < 0.001$

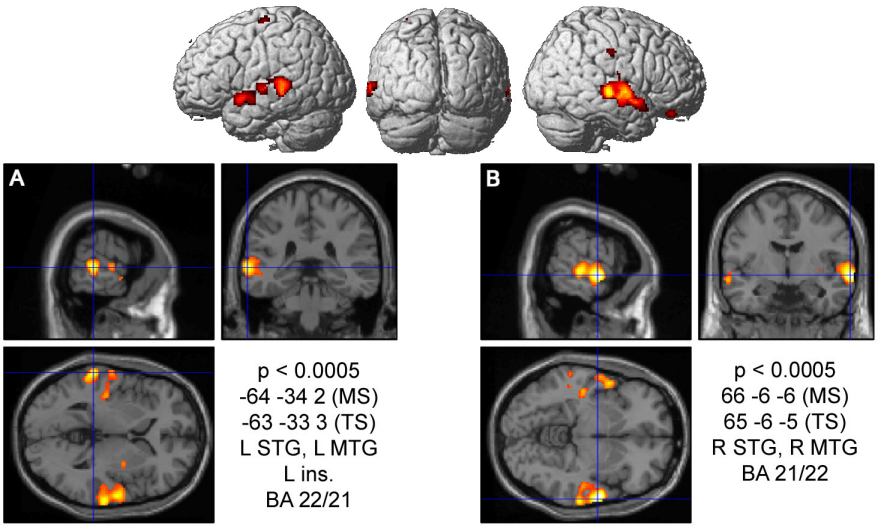


FIG. 13.14 – Activations cérébrales de groupe observées pour les patients implantés testés à T1 (n = 8). Comparaison entre les activations par de la parole audiovisuelle et visuelle (contraste AV-V). Volumes supérieurs à 30 voxels avec $p < 0.001$

13.2.4 Evolution longitudinale des activations

Suractivations à T0

En condition visuelle (V-R), les patients implantés testés peu après la mise en marche de l'implant (T0) présentent une activation significativement plus élevée qu'à T1 dans la partie la plus antérieure du lobe temporal droit, notamment dans les régions cérébrales sélectives à l'identité de la voix (fig. 13.15). En condition audiovisuelle, on observe une suractivation thalamique dans le noyau médio-dorsal bilatéral et ventro-latéral gauche (fig. 13.16).

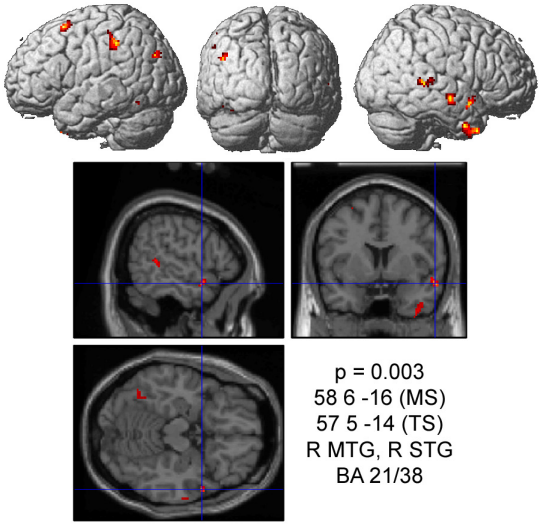


FIG. 13.15 – Comparaison appariée des activations cérébrales des patients implantés testés à T0 et T1 (n = 8 ; T0 > T1). Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

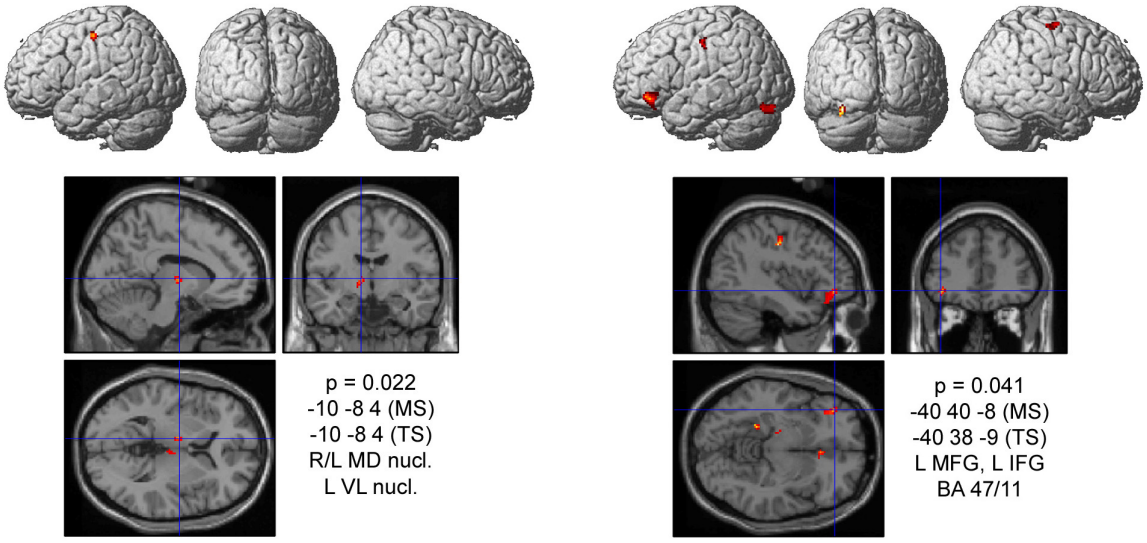


FIG. 13.16 – Comparaison appariée des activations cérébrales des patients implantés testés à T0 et T1 (n = 8 ; T0 > T1). Parole audiovisuelle (contrastes AV-R et AV-V). Volumes supérieurs à 30 voxels avec $p < 0.005$

Suractivations à T1

En condition visuelle, les patients implantés testés après récupération perceptive (T1) présentent des activations significativement plus élevées qu'à T0 dans le cortex frontal gauche. On observe notamment une suractivation dans l'aire de Broca en condition de référence motrice (B, fig. 13.17, p. 306), et une suractivation dans le cortex frontal inférieur ventral gauche (fig. 13.18). En condition de référence motrice, on peut également remarquer une suractivation à T1 dans le cortex frontal postérieur dorsal droit (A, fig. 13.17).

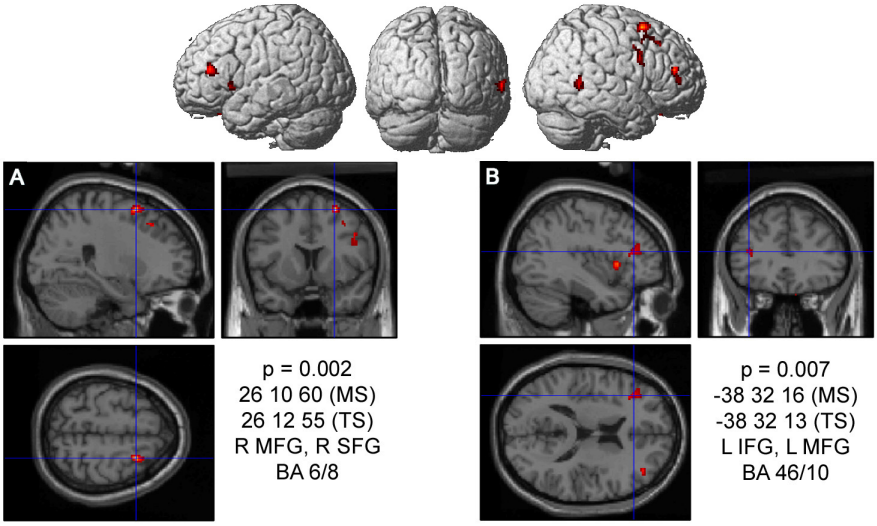


FIG. 13.17 – Comparaison appariée des activations cérébrales des patients implantés testés à T0 et T1 (n = 8 ; T1 > T0). Condition de référence motrice (contraste RM-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

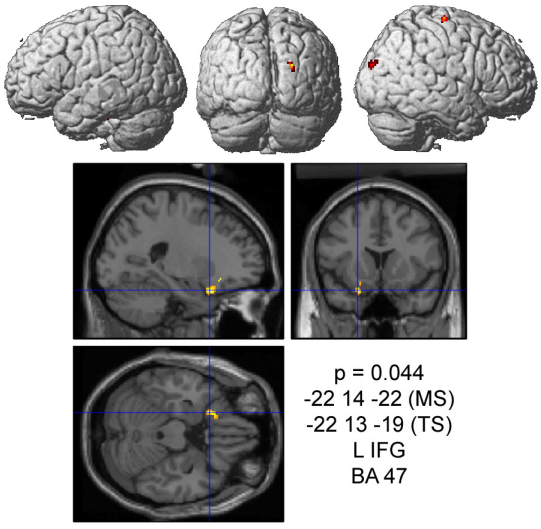


FIG. 13.18 – Comparaison appariée des activations cérébrales des patients implantés testés à T0 et T1 (n = 8 ; T1 > T0). Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

13.2.5 Récapitulatif des activations cérébrales

Les activations observées chez les implantés testés peu après la mise en marche de l'implant (T0) puis après récupération perceptive (T1) tendent à démontrer des phénomènes de réorganisation corticale accompagnés d'un changement de latéralisation du traitement de la parole. En effet, les activations spécifiques au groupe T0 sont pour une grande majorité latéralisées dans l'hémisphère droit, en particulier au niveau des aires auditives primaires et associatives réparties le long du gyrus temporal supérieur, ainsi qu'au niveau du cortex frontal inférieur antérieur droit. Au contraire, les activations spécifiques au groupe T1 sont pour une grande majorité latéralisées dans l'hémisphère gauche, en particulier au niveau des régions corticales frontales moyennes et inférieures avoisinant l'aire de Broca, ainsi qu'au niveau du cortex auditif associatif gauche.

Contraste	**Activations différenciant les groupes T0 et T1**	
	T0	**T1**
RM-R	temporal inférieur antérieur D	*frontal moyen latéral G* *frontal supérieur dorsal D*
V-R	*temporal supérieur antérieur D* temporal supérieur postérieur D	*frontal inférieur ventral G*
V-RM	frontal inférieur antérieur D cervelet latéral D	frontal moyen antérieur G
AV-RM	auditif primaire/associatif D (BA 41/42) frontal inférieur antérieur D précentral ventral D	auditif associatif G (BA 38)
AV-V	auditif primaire/associatif D (BA 42) *frontal inférieur postérieur G*	

TAB. 13.3 – Principales activations significatives différenciant les patients implantés testés à T0 et à T1. En italiques, activations apparaissant comme significatives lors de la comparaison appariée entre les deux groupes.

Contraste	**Activations en commun entre les groupes T0 et T1**
RM-R	occipital moyen et inférieur bilatéral fusiforme/lingual/cuneus bilatéral postcentral/lobule pariétal inférieur D
V-R	occipital moyen et inférieur bilatéral fusiforme/lingual/déclive bilatéral postcentral/lobule pariétal inférieur D temporal supérieur postérieur G
AV-RM	auditif associatif bilatéral (BA 21/22)
AV-V	auditif associatif bilatéral (BA 21/22)

TAB. 13.4 – Principales activations significatives en commun entre les patients implantés testés à T0 et à T1.

13.3 Activations cérébrales chez les sujets contrôles

13.3.1 Activations visuelles

Contraste RM-R

Les activations cérébrales observées dans notre groupe de 6 sujets contrôles normo-entendants lors de la présentation de visages en contexte non-linguistique font principalement intervenir le réseau bilatéral des aires visuelles associatives (voir fig. 13.19, p. 308). On observe ainsi des foyers d'activation dans la partie postérieure des gyri occipitaux moyen et inférieur, ainsi que dans le cuneus et les gyri linguaux et fusiformes (A, B et D, fig. 13.19, p. 308). On retrouve en outre l'activation motrice induite par la réponse manuelle à l'aide de la souris d'ordinateur, se traduisant par un foyer d'activation localisé entre le lobule pariétal inférieur gauche et la partie dorso-latérale du gyrus post-central gauche (C, fig. 13.19, p. 308).

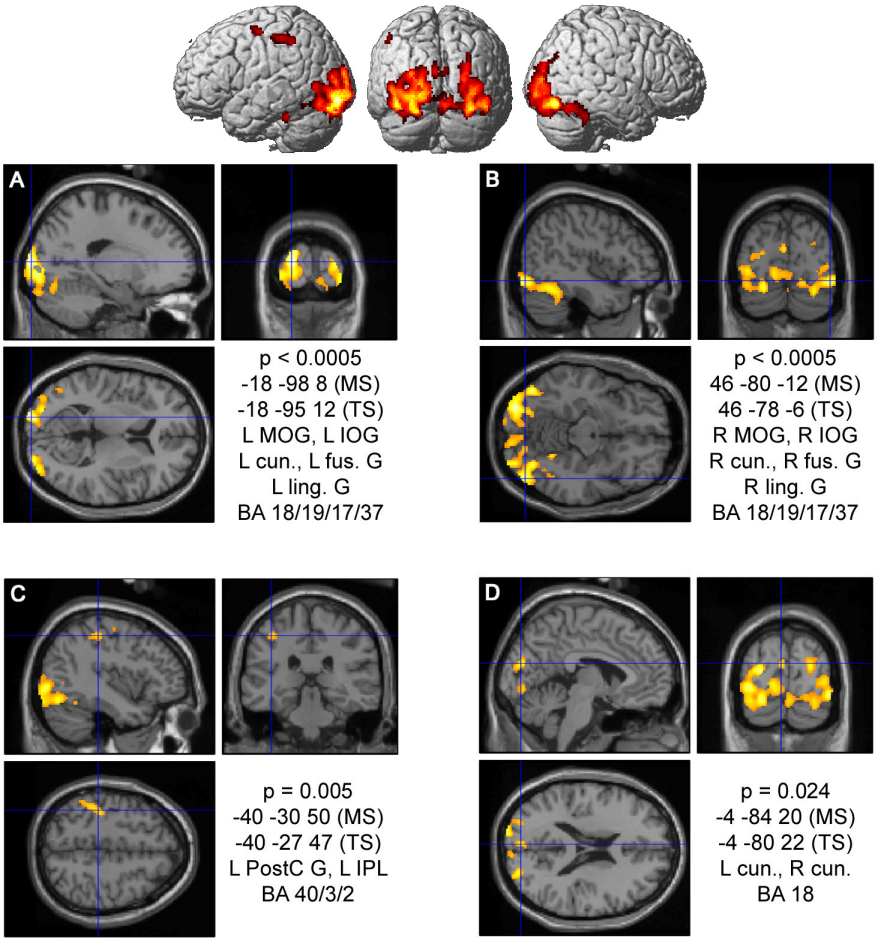


FIG. 13.19 – Activations cérébrales de groupe observées pour l'ensemble des 6 sujets contrôles. Présentation des visages sans contexte linguistique (contraste RM-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

Contraste V-R

Les activations induites dans notre groupe de 6 sujets normo-entendants par la présentation de parole visuelle présentent quelques similarités avec celles observées pour la même condition chez les patients implantés cochléaires (voir fig. 13.20, p. 309). Ainsi on retrouve le réseau visuel associatif formé des parties postérieures bilatérales des gyri fusiformes et des gyri occipitaux moyens et inférieurs, ainsi que les gyri linguaux et le cuneus (A, B, C et D, fig. 13.20, p. 309).

A la différence des patients implantés testés à T0, on ne retrouve d'activation ni dans le sillon temporal supérieur postérieur, ni dans le lobe temporal antérieur droit, ni dans le lobe frontal inférieur droit. Les sujets normo-entendants présentent au contraire des foyers d'activation dans le lobe frontal inférieur et moyen gauche, au niveau de l'aire de Broca (E et F), ces activations n'ayant pas été observées dans le groupe de patients implantés testés à T0.

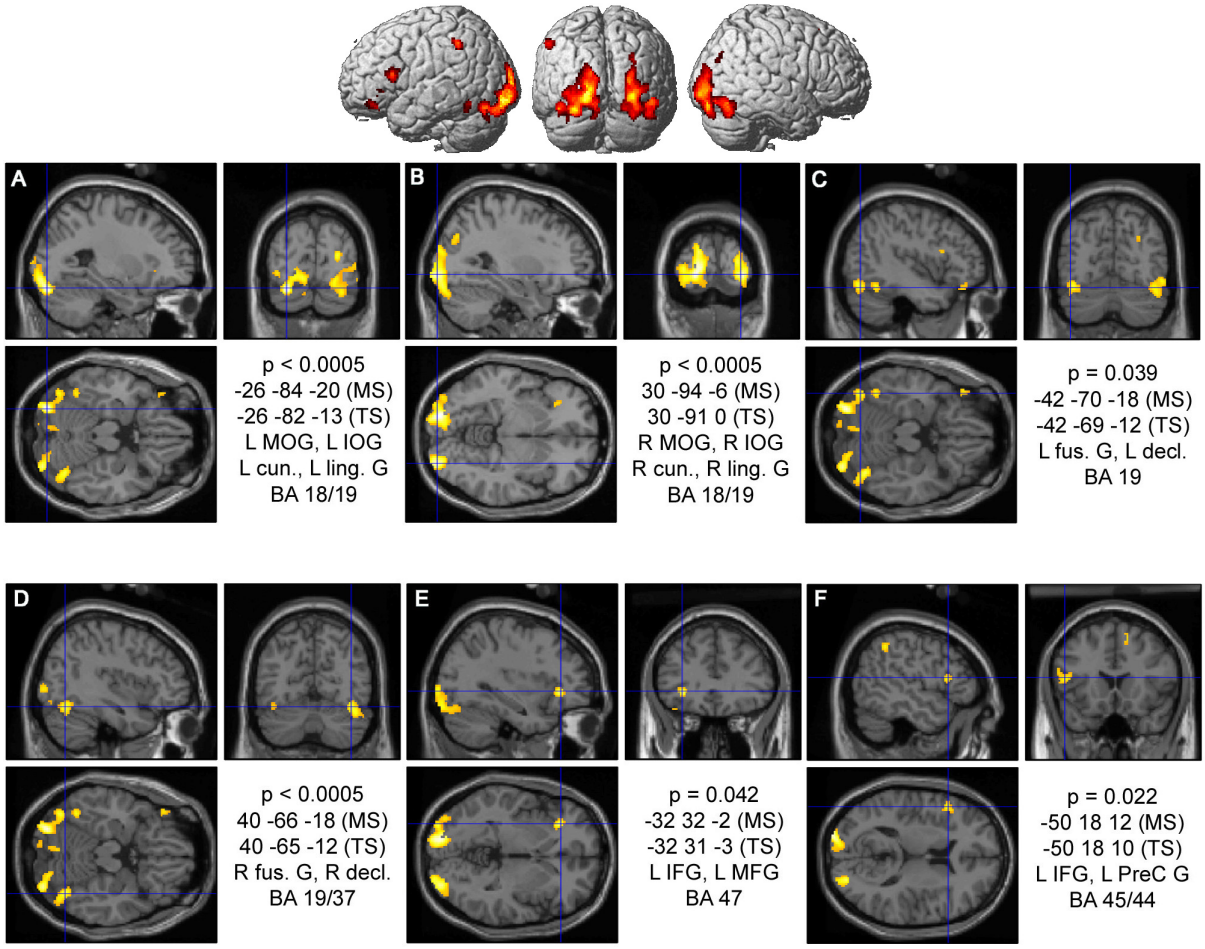


FIG. 13.20 – Activations cérébrales de groupe observées pour l'ensemble des 6 sujets contrôles. Parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.001$

Contraste V-RM

Contrairement aux patients implantés cochléaires, les sujets normo-entendants ne présentent aucun foyer d'activation statistiquement significatif (au seuil $p = 0.05$) lors de la tâche de perception de la parole visuelle contrastée avec la présentation non-linguistique des visages.

13.3.2 Activations audiovisuelles

Contraste AV-RM

Lors de la présentation de parole audiovisuelle (contrastée avec la tâche de présentation non-linguistique des visages), les sujets normo-entendants présentent des activations des aires auditives associatives, de façon similaire aux patients implantés cochléaires (A et B, fig. 13.21, p. 310). Toutefois, alors que les patients implantés présentaient une légère asymétrie en faveur de l'hémisphère droit, les sujets normo-entendants présentent ici une très forte asymétrie en faveur de l'hémisphère gauche, conformément aux données classiquement rapportées dans la littérature scientifique (voir *« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91).

Cette asymétrie en faveur de l'hémisphère gauche s'étend en particulier dans le gyrus temporal transverse gauche (A) ainsi que dans le lobe frontal inférieur et moyen gauche, au niveau de l'aire de Broca (C et D, fig. 13.21, p. 310). Pour la même condition, ces activations du lobe frontal gauche ne sont pas observées chez les patients implantés, tandis que ces derniers présentent des activations du lobe frontal droit qui ne sont pas présentes chez les sujets normo-entendants.

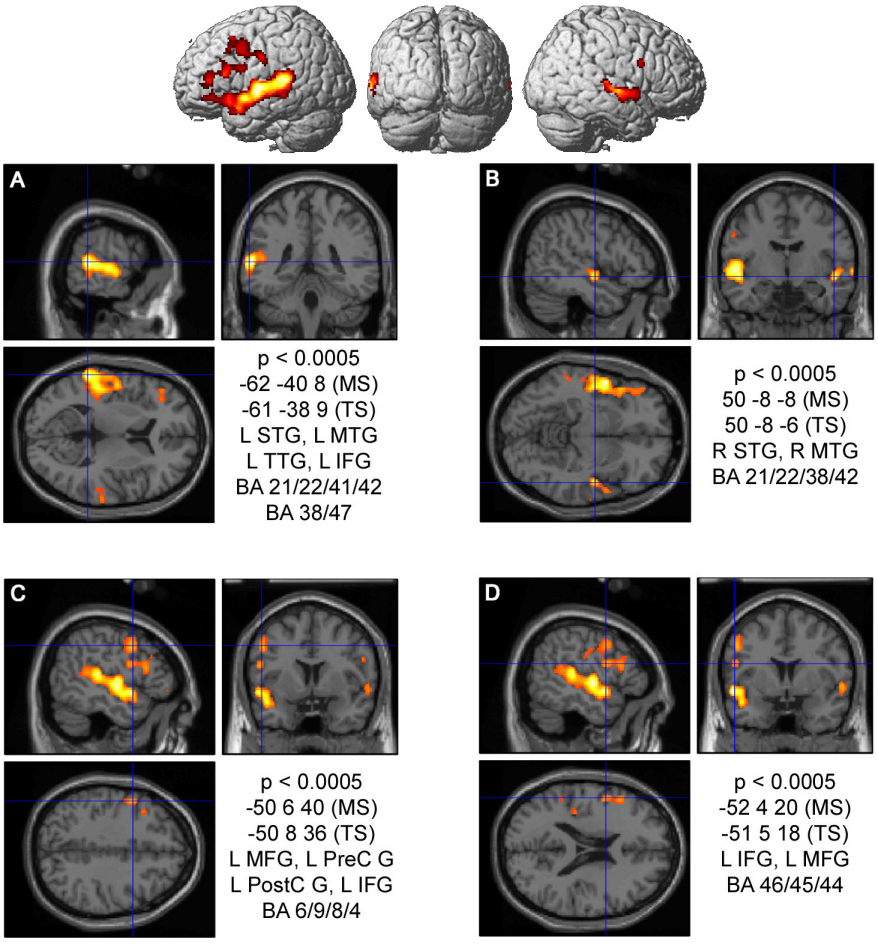


FIG. 13.21 – Activations cérébrales de groupe observées pour l'ensemble des 6 sujets contrôles. Parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.001$

Contraste AV-V

La comparaison des conditions de parole audiovisuelle et de parole visuelle chez les sujets normo-entendants indique des suractivations des aires auditives associatives bilatérales (aires de Brodmann 21, 22 et 41) (voir fig. 13.22, p. 311). Contrairement à ce qui avait été observé dans le cas des patients implantés cochléaires, ces suractivations présentent une asymétrie en faveur de l'hémisphère gauche, alors que l'asymétrie observée chez les patients implantés était en faveur de l'hémisphère droit.

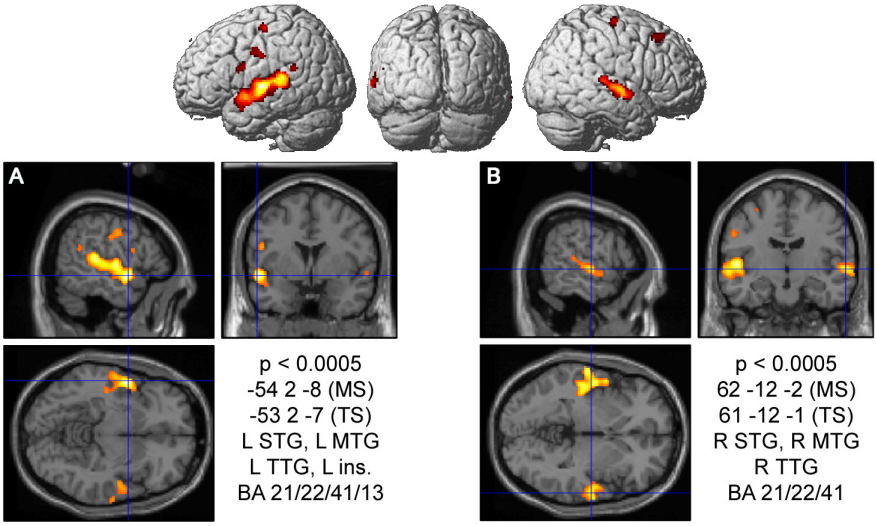


FIG. 13.22 – Activations cérébrales de groupe observées pour l'ensemble des 6 sujets contrôles. Comparaison entre les activations par de la parole audiovisuelle et visuelle (contraste AV-V). Volumes supérieurs à 30 voxels avec $p < 0.001$

13.3.3 Récapitulatif des activations cérébrales

Contraste	Activations de groupe
RM-R	occipital moyen et inférieur bilatéral fusiforme/lingual/cuneus bilatéral postcentral/lobule pariétal inférieur D
V-R	occipital moyen et inférieur bilatéral fusiforme/lingual/déclive bilatéral postcentral/lobule pariétal inférieur D frontal inférieur postérieur G
AV-RM	auditif associatif bilatéral (BA 21/22) auditif primaire gauche (BA 41/42) frontal inférieur postérieur G
AV-V	auditif primaire/associatif bilatéral (BA 21/22/41)

TAB. 13.5 – Principales activations significatives observées dans le groupe de contrôles normo-entendants.

13.4 Différences d'activité patients/contrôles

13.4.1 Différences d'activité en stimulation visuelle

Contraste RM-R

Lors d'une tâche de visualisation des visages sans contexte linguistique, on n'observe aucune différence significative entre les activations observées chez les sujets contrôles et les patients implantés, que ceux-ci soient testés à la mise en marche de l'implant (T0) ou après récupération perceptive (T1).

Contraste V-R

Lors d'une tâche de lecture labiale contrastée avec la condition de repos (V-R), les sujets contrôles présentent une suractivation significative spécifiquement située dans l'aire de Broca (partie postérieure gauche des gyri frontaux inférieur et moyen ; voir fig. 13.23, p. 312). Cette suractivation est présente pour la comparaison avec les patients implantés testés précocément (T0) mais disparaît pour la comparaison avec les patients implantés testés tardivement (T1).

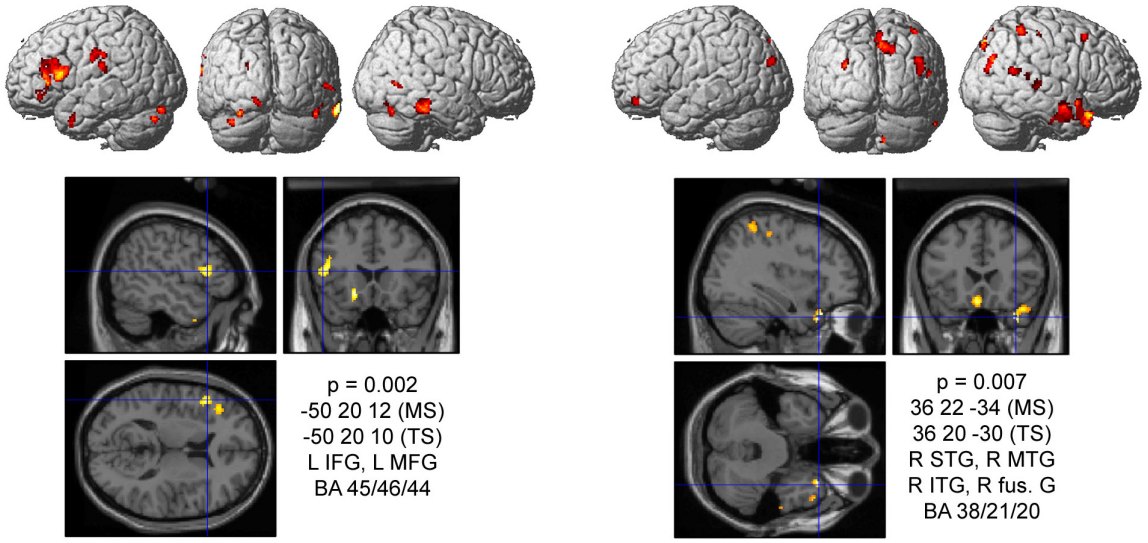


FIG. 13.23 – Différences NH (n = 6) > IC à T0 (n = 10) et IC à T0 (n = 10) > NH (n = 6) en condition de parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

Les patients implantés testés à T0 présentent quant à eux une suractivation significative de la partie la plus antérieure du lobe temporal droit (partie antérieure droite des gyri temporaux supérieur et moyen ; voir fig. 13.23, p. 312). Cette suractivation disparaît à T1, remplacée par une suractivation de la partie antérieure du gyrus frontal moyen droit (voir fig. 13.24, p. 313)

Contraste V-RM

Lors de la tâche de lecture labiale contrastée avec la tâche de visualisation des visages sans contexte linguistique (V-RM), on n'observe aucune différence significative entre les activations observées chez les sujets contrôles et les patients implantés, que ceux-ci soient testés à la mise en marche de l'implant (T0) ou après récupération perceptive (T1).

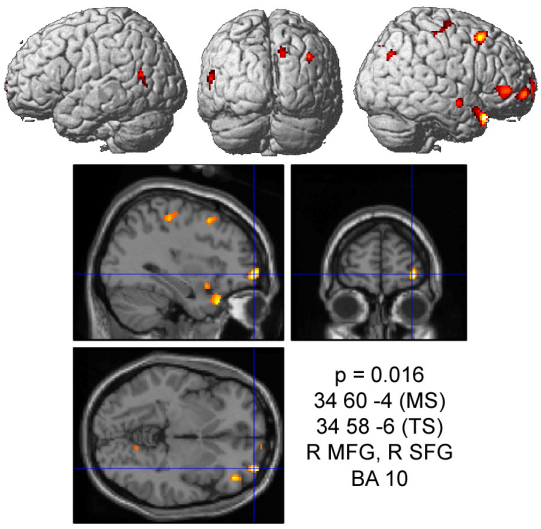


FIG. 13.24 – Différences IC à T1 (n = 8) > NH (n = 6) en condition de parole visuelle (contraste V-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

13.4.2 Différences d'activité en stimulation audiovisuelle

Contraste AV-R

Lors de la tâche de parole audiovisuelle contrastée avec la condition de repos, on peut remarquer chez les sujets normo-entendants plusieurs foyers de suractivation concentrés dans l'hémisphère gauche, ces suractivations existant à la fois dans les comparaisons avec les implantés testés précocément (T0) et tardivement (T1). Parmi ces foyers, on distingue l'aire de Broca (partie postérieure gauche des gyri frontaux inférieur et moyen, voir A, fig. 13.25, p. 314 et fig. 13.26, p. 314), la région suractivée correspondant à une extension de celle observée dans la condition V-R (voir fig. 13.23, p. 312).

On distingue en outre un foyer de suractivation dans l'aire prémotrice (partie moyenne du gyrus précentral gauche, voir C, fig. 13.25 et 13.26), correspondant probablement à la zone codant pour les mouvements de la bouche (cette aire semblait également suractivée dans la condition V-R, sans toutefois dépasser le seuil de significativité statistique). On note par ailleurs un foyer de suractivation touchant les régions auditives gauches primaire et secondaire (gyrus temporal transverse gauche et partie dorsale du gyrus temporal supérieur gauche, dans le prolongement du gyrus précentral, voir B, fig. 13.25 et 13.26). Enfin, on observe également une suractivation du gyrus cingulaire, dans sa partie la plus dorsale (D), toutefois cette suractivation n'apparaît que dans la comparaison avec les patients testés à la mise en marche de l'implant (T0, voir fig. 13.25, p. 314).

Inversement, les suractivations présentées par les patients implantés cochléaires sont essentiellement concentrées dans l'hémisphère droit. Parmi les aires concernées pour les patients testés à T0, on remarque une suractivation de la partie la plus antérieure du lobe temporal droit (partie antérieure droite des gyri temporaux supérieur et moyen, voir A, fig. 13.27, p. 315), cette aire correspondant à une extension de la région observée dans la condition V-R (voir fig. 13.23, p. 312).

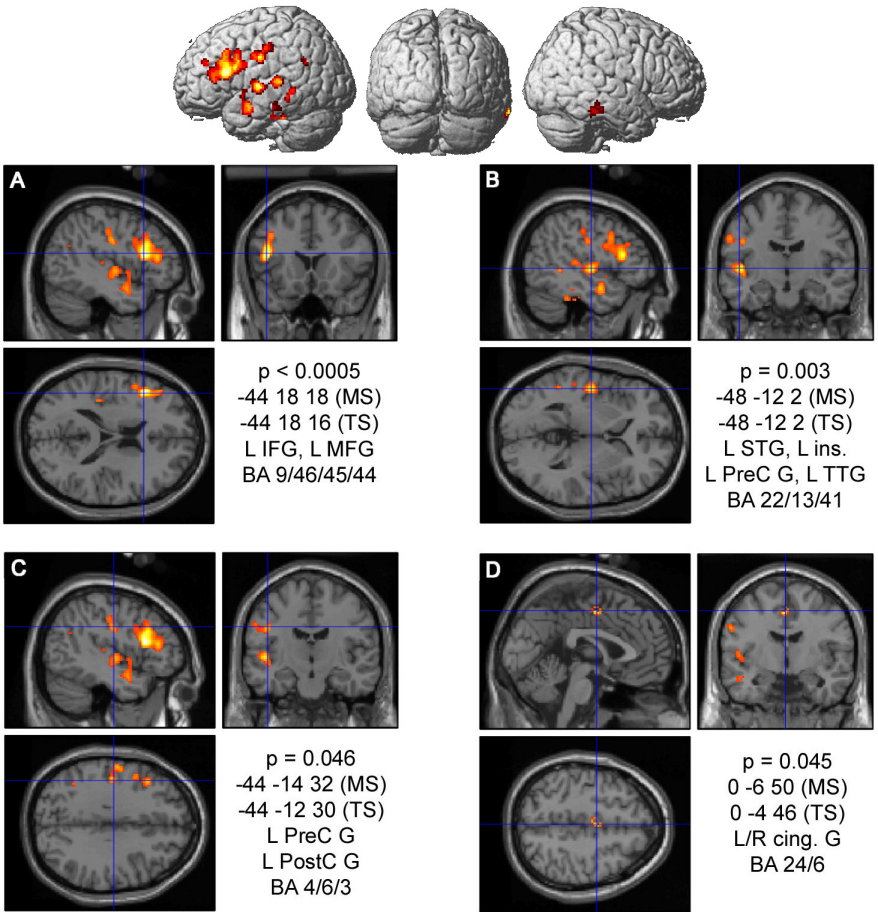


FIG. 13.25 – Différences NH (n = 6) > IC à T0 (n = 10) en condition de parole audiovisuelle (contraste AV-R). Volumes supérieurs à 30 voxels avec p < 0.005

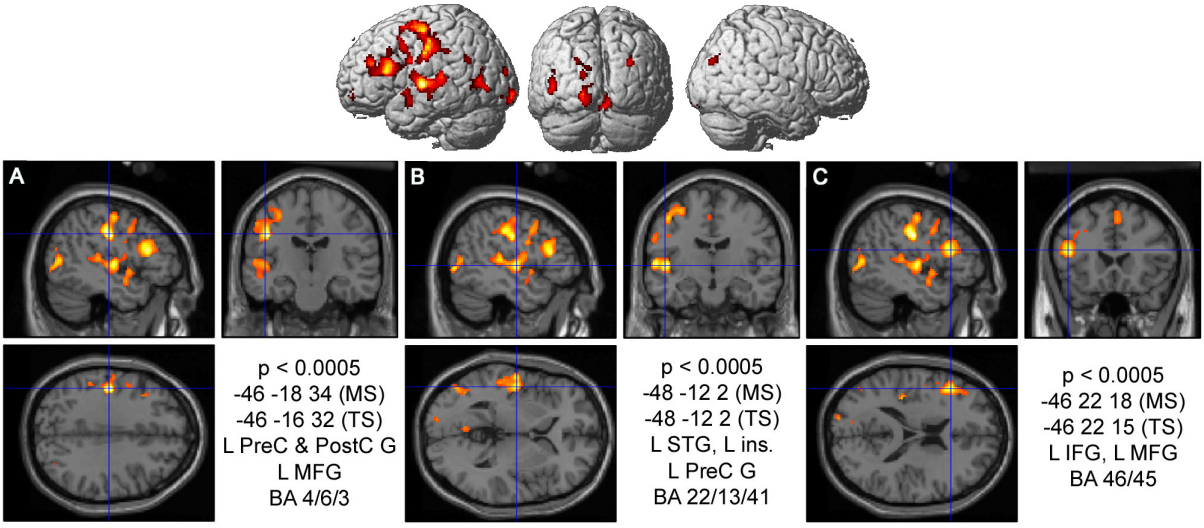


FIG. 13.26 – Différences NH (n = 6) > IC à T1 (n = 8) en condition de parole audiovisuelle (contraste AV-R). Volumes supérieurs à 30 voxels avec p < 0.005

On observe également une suractivation de la jonction temporo-pariéto-occipitale droite, touchant à la fois le gyrus supramarginal, la partie postérieure du gyrus temporal supérieur et le lobule pariétal inférieur (voir C, fig. 13.27, p. 315 ; cette aire semblait également suractivée dans la condition V-R, sans toutefois dépasser le seuil de significativité statistique). Pour terminer, on observe également un foyer de suractivation dans la partie ventrale médiale postérieure du lobe frontal, avec une asymétrie en faveur de l'hémisphère gauche, entre les aires de Brodmann 25 et 11 (B, fig. 13.27, p. 315). Pour les patients implantés testés après récupération perceptive (à T1), on observe une disparition des ces foyers de suractivation. On observe également l'apparition d'un foyer de suractivation dans les profondeurs de la partie antérieure du sillon temporal supérieur droit (fig. 13.28, p. 315).

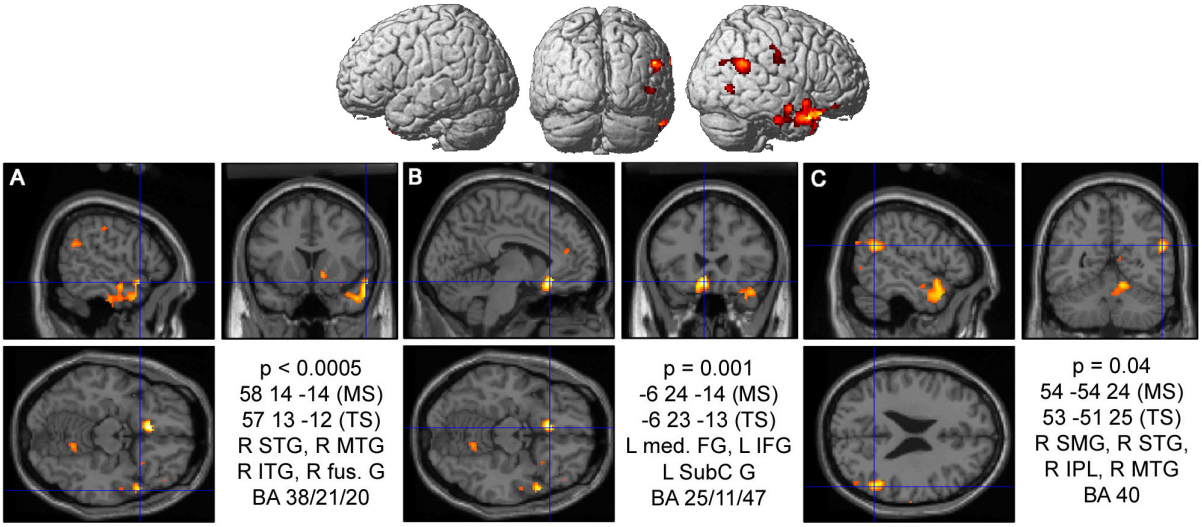


FIG. 13.27 – Différences IC à T0 (n = 10) > NH (n = 6) en condition de parole audiovisuelle (contraste AV-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

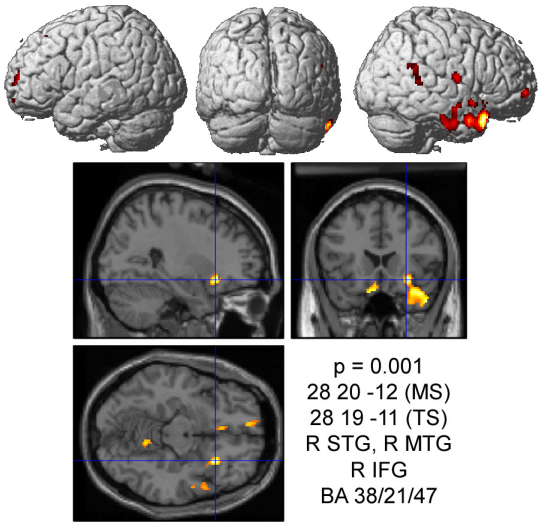


FIG. 13.28 – Différences IC à T1 (n = 8) > NH (n = 6) en condition de parole audiovisuelle (contraste AV-R). Volumes supérieurs à 30 voxels avec $p < 0.005$

Contraste AV-RM

Lors de la tâche de parole audiovisuelle contrastée avec la visualisation des visages sans contexte linguistique, les patients implantés testés à T0 présentent une sous-activation significative du cortex auditif primaire et associatif gauche (gyrus temporal transverse gauche et partie dorsale du gyrus temporal supérieur gauche ; voir fig. 13.29, p. 316). Les patients implantés testés à T1 présentent également des sous-activations similaires (voir fig. 13.30, p. 316).

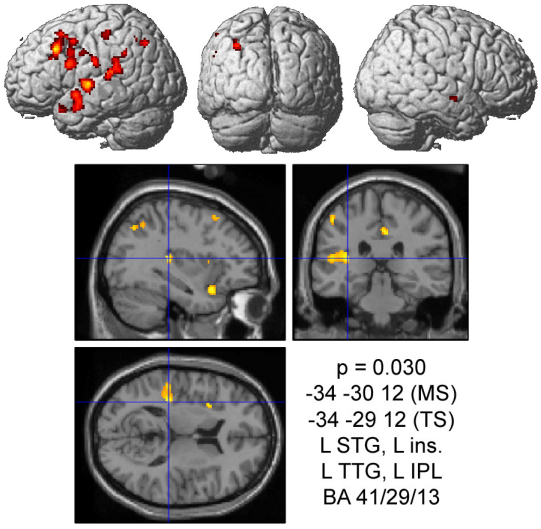


FIG. 13.29 – Différences NH (n = 6) > IC à T0 (n = 10) en condition de parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

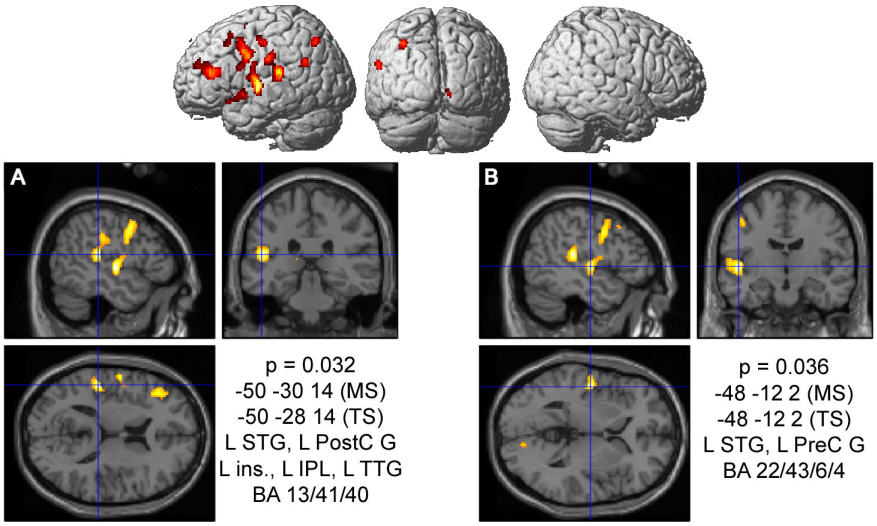


FIG. 13.30 – Différences NH (n = 6) > IC à T1 (n = 8) en condition de parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

Au niveau des suractivations, les patients implantés testés à T0 présentent une suractiva-

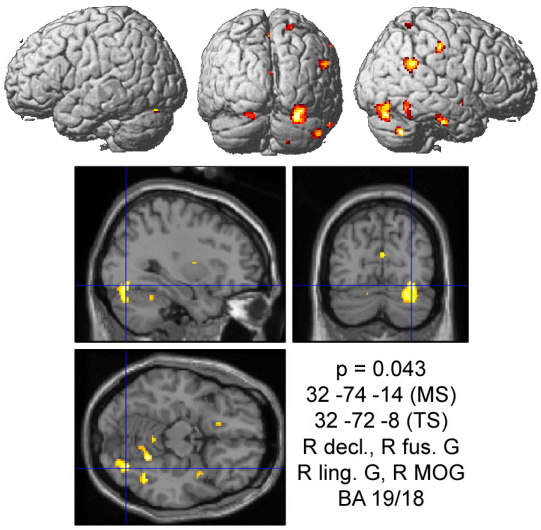


FIG. 13.31 – Différences IC à T0 (n = 10) > NH (n = 6) en condition de parole audiovisuelle avec activations spécifiques du langage (contraste AV-RM). Volumes supérieurs à 30 voxels avec $p < 0.005$

tion du cortex visuel associatif droit (partie postérieure du gyrus fusiforme droit ; voir fig. 13.31, p. 317) qui disparaît après récupération perceptive, n'apparaissant plus à T1.

Contraste AV-V

Dans le cas de la tâche de parole audiovisuelle contrastée avec la tâche de lecture labiale, les patients implantés présentent une sous-activation du cortex auditif primaire et associatif gauche similaire à celle qui avait pu être observée pour les contrastes AV-RM et AV-R (voir fig. 13.32, p. 317).

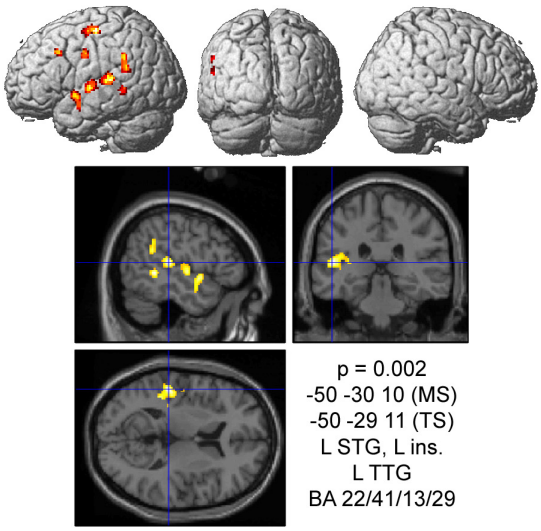


FIG. 13.32 – Différences NH (n = 6) > IC à T0 (n = 10) pour la comparaison entre les activations par de la parole audiovisuelle et visuelle (contraste AV-V). Volumes supérieurs à 30 voxels avec $p < 0.005$

Cette sous-activation persiste même après récupération perceptive (voir A, fig. 13.33, p. 318). À T1, on voit en outre apparaître une sous-activation de la partie dorso-latérale des régions motrices gauches, correspondant probablement à la partie codant pour les mouvements de la bouche (voir B, fig. 13.33).

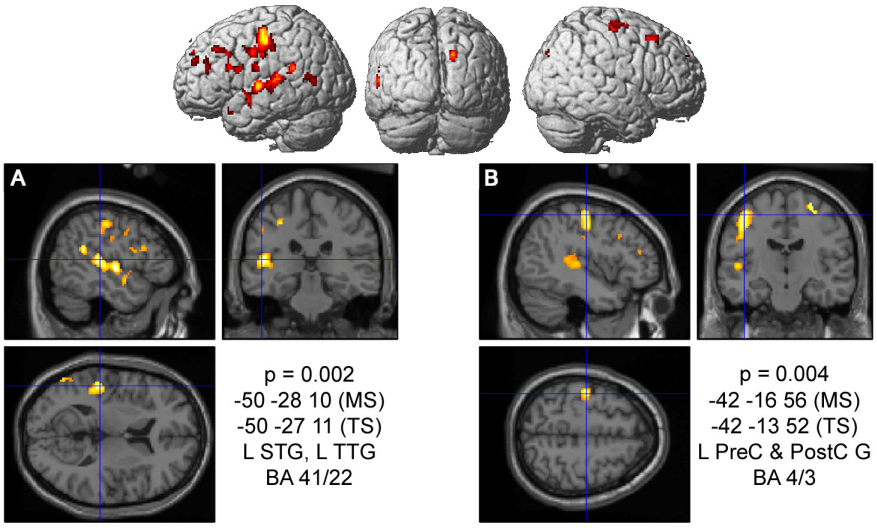


FIG. 13.33 – Différences NH (n = 6) > IC à T1 (n = 8) pour la comparaison entre les activations par de la parole audiovisuelle et visuelle (contraste AV-V). Volumes supérieurs à 30 voxels avec $p < 0.005$

13.4.3 Récapitulatif des différences d'activations cérébrales

La comparaison des activations cérébrales observées chez les patients implantés et les contrôles normo-entendants démontre une très nette différence de latéralisation entre les deux groupes. Ainsi les suractivations présentées par les sujets contrôles sont systématiquement latéralisées dans l'hémisphère gauche, en particulier dans l'aire de Broca et dans le cortex auditif primaire et associatif gauche, tandis que les suractivations présentées par les patients implantés sont pour la plupart latéralisées dans l'hémisphère droit, en particulier dans le gyrus temporal supérieur antérieur droit et dans le cortex frontal droit.

Contraste	**Activations différenciant contrôles et implantés testés à T0**	
	NH	**IC à T0**
V-R	frontal inférieur postérieur G	temporal supérieur antérieur D
AV-R	frontal inférieur postérieur G précentral latéral G auditif primaire/associatif G cingulaire antérieur médial	temporal supérieur antérieur D supramarginal D frontal ventral médial G
AV-RM	auditif primaire/associatif G	fusiforme D
AV-V	auditif primaire/associatif G	

TAB. 13.6 – Différences significatives d'activations différenciant le groupe de contrôles normo-entendants et le groupe des patients implantés testés à T0.

Contraste	Activations différenciant contrôles et implantés testés à T1	
	NH	IC à T1
V-R		frontal moyen antérieur D
AV-R	frontal inférieur postérieur G précentral latéral G auditif associatif G	temporal supérieur antérieur D
AV-RM	auditif primaire/associatif G	
AV-V	auditif primaire/associatif G	

TAB. 13.7 – Différences significatives d'activations différenciant le groupe de contrôles normo-entendants et le groupe des patients implantés testés à T1.

13.5 Etude de corrélation des activations cérébrales

13.5.1 Méthodes

Afin d'inclure nos données tomographiques dans une analyse de corrélation fonctionnelle, nous avons isolé les pics d'activation suivants (en coordonnées MNI) : aire frontale de Broca ([-50, 20, 12]), aire temporale de la voix ([58, 12, -14]), gyrus frontal moyen droit ([34, 60, -4]), gyrus supramarginal droit ([54, -54, 24]), cortex auditif gauche ([-48, -12, 2] et [-50, -30, 10]), cortex temporal supérieur postérieur droit ([62, -38, 6]) et gauche ([-56, -46, 10]), aire fusiforme du visage droite ([34, -62, -18]) et gauche ([-34, -54, -24]). Pour chacun des ces pics d'activation, nous avons relevé le début sanguin cérébral dans une région sphérique s'étendant jusqu'à 4 mm autour du pic, puis utilisé ces mesures de débit dans notre analyse de corrélation fonctionnelle.

13.5.2 Corrélation entre données comportementales et activations cérébrales

Nous avons examiné la présence d'une éventuelle corrélation entre les scores comportementaux obtenus en lecture labiale et les débits sanguins cérébraux régionaux. L'analyse statistique des résultats comportementaux obtenus à T0 et T1 ne révèle qu'une seule corrélation significative avec les activations observées : en effet on observe que le degré d'activité à T0 (après la mise en marche de l'implant) dans l'aire temporale de la voix (partie antérieure du sulcus temporal supérieur droit) est positivement corrélé aux scores d'identification en lecture labiale, que ceux-ci soient issus de l'évaluation orthophonique avant la tomographie ($p = 0.0196$) ou obtenus directement pendant la séance de tomographie ($p = 0.0008$).

13.5.3 Connectivité cérébrale

Au travers d'une analyse de la connectivité fonctionnelle cérébrale, nous avons pu mettre en évidence les corrélations suivantes. L'activité dans l'aire de Broca pendant la tâche de lecture labiale est corrélée de façon positive avec l'activité dans l'aire temporale supérieure postérieure gauche (proche de la jonction temporo-pariéto-occipitale) à T0 ($p = 0.039$) ainsi qu'à T1 ($p = 0.050$), mais pas dans le groupe contrôle ($p = 0.542$). L'activité dans l'aire de Broca est également corrélée avec l'activité dans le cortex auditif gauche ([-48, -12, 2]) à T0 ($p = 0.025$) et T1 ($p = 0.039$), mais pas dans le groupe contrôle malgré une légère tendance ($p = 0.094$).

Enfin, de façon particulièrement intéressante, l'activité dans l'aire de Broca semble n'être corrélée avec l'activité dans l'aire temporale de la voix (partie antérieure du sulcus temporal supérieur droit) pour aucun des trois groupes ($p = 0.663$, 0.862 et 0.339 respectivement). Toutefois

une très nette corrélation négative apparaît entre les activités dans l'aire de Broca et l'aire de la voix lorsque les trois groupes sont rassemblés ($p = 0.009$).

13.6 Discussion

13.6.1 Réorganisation cérébrale du traitement visuel de la parole

Pendant les vingt dernières années, de nombreuses études en neuroimagerie ont été menées en vue de circonscrire le réseau cortical impliqué dans la compréhension de la parole visuelle chez le sujet normo-entendant et les patients sourds implantés cochléaires (Campbell et Capek, 2008). La plupart de ces études s'accordent sur un chevauchement étendu des régions cérébrales activées par la lecture labiale dans ces deux catégories de sujets. Ainsi, pour le groupe de patients implantés comme pour le groupe de sujets contrôles, notre tâche de lecture labiale induit des activations du réseau cortical spécialisé dans l'analyse visuelle statique et dynamique du visage. Ce réseau inclut les aires visuelles associatives dont en particulier le gyrus fusiforme bilatéral, activé lors de la reconnaissance visuelle des visages (Kanwisher *et al.*, 1997; McCarthy *et al.*, 1997; Haxby *et al.*, 2000), ainsi que l'aire occipitale des visages, située sur la partie latérale du gyrus occipital inférieur et également impliquée dans le traitement visuel des visages (Kanwisher et Yovel, 2006) (voir *« Substrats neuronaux de la parole visuelle »*, sect. 4.3, p. 82).

Toutefois certaines différences peuvent être détectées qui sont principalement limitées aux aires auditives temporales et au cortex frontal. Dans notre étude de discrimination entre mots et non-mots, nous démontrons des disparités dans l'activité cérébrale entre les sujets normo-entendants et les patients sourds implantés cochléaires. Ces disparités incluent des régions frontales et temporales, en accord avec les données de la littérature (Lee *et al.*, 2007; Capek *et al.*, 2008).

Réorganisation dans les régions temporales

Pendant la tâche de lecture labiale, nous avons observé une activation des aires auditives temporales (BA 22) chez les patients sourds, indépendamment de la durée d'expérience avec l'implant. Toutefois, l'activation des régions auditives par la parole visuelle ne s'est pas étendue au cortex auditif primaire. Cela confirme des études précédentes réalisées chez des patients implantés qui ont reporté une activation du cortex auditif par le langage signé visuel (Nishimura *et al.*, 1999; Petitto et Kovelman, 2003; Sadato *et al.*, 2004; Lambertz *et al.*, 2005) ou par la lecture labiale visuelle (Lee *et al.*, 2007; Capek *et al.*, 2008; Campbell et Capek, 2008), ces activations excluant le cortex auditif primaire. Bien que des activations des aires auditives par la lecture labiale aient déjà été reportées chez des sujets normo-entendants (Bernstein *et al.*, 2002; Paulesu *et al.*, 2003), nous n'avons pas pu reproduire de telles activations ; cette disparité pourrait être liée à des différences dans les tâches de lecture labiale dans lesquelles les sujets sont impliqués.

De façon particulièrement intéressante, les patients implantés cochléaires présentent une suractivation significative de la partie la plus antérieure du gyrus temporal supérieur droit pendant la lecture labiale. Cette partie du cortex, l'« aire temporale de la voix » est une région auditive associative spécifiquement impliquée lors de la reconnaissance de l'identité de la voix du locuteur (Belin et Zatorre, 2003; Fecteau *et al.*, 2004a,b; von Kriegstein et Giraud, 2004) (voir *« Reconnaissance de la voix humaine »*, § 4.2.1, p. 81). Comme cela sera discuté en détail dans la suite de cette discussion, l'aire de la voix est susceptible de subir une réorganisation intermodale pendant la surdité, cette restructuration pouvant se traduire par une participation au traitement visuel de la parole.

Une différence essentielle qui apparaît entre les patients implantés et les sujets normo-entendants consiste dans la suractivation bilatérale de la partie postérieure du sillon temporal supérieur chez les patients implantés, au niveau du gyrus angulaire, tandis que cette zone reste inactivée chez les sujets contrôles. Cette partie du cortex est connue en particulier pour intervenir dans l'intégration phonologique de la parole (Hickok et Poeppel, 2004; Démonet *et al.*, 2005; Skipper *et al.*, 2005; Vigneau *et al.*, 2006) (voir *« Intégration corticale audio-visuo-motrice de la parole »*, § 4.4.3, p. 86) mais également dans l'association des mots écrits avec leur correspondance auditive (Démonet *et al.*, 1992, 2005). Ainsi, l'expertise comportementale en lecture labiale observée chez les patients implantés cochléaires pourrait se traduire au niveau neurofonctionnel par une suractivation des réseaux spécifiquement impliqués dans l'intégration phonologique de la parole. De cette façon, la vision (réelle) des mouvements des lèvres serait associée à l'audition (virtuelle) des mots associés grâce à un phénomène de coopération intermodale qui faciliterait l'induction d'une représentation phonologique à partir du percept audiovisuel.

Enfin, dans le cadre des activations intermodales des aires temporales par la parole visuelle, il est intéressant d'observer l'apparition chez les patients implantés testés à T1 en condition de parole visuelle (V-RM) d'une zone localement suractivée dans la partie moyenne du sillon temporal supérieur gauche (voir fig. 13.10, p. 301), la même région ne présentant aucune suractivation à T0. Toutefois cette suractivation se maintient un peu en-dessous du seuil de significativité statistique. En considérant le nombre restreint de patients dans le groupe testé à T1, il n'est pas impossible que cette suractivation, bien que non significative, soit révélatrice d'un phénomène de coopération intermodale entre l'audition et la vision se mettant progressivement en place lors de la récupération fonctionnelle auditive des patients implantés cochléaires.

Réorganisation dans les régions frontales

Chez les contrôles normo-entendants, des activations de l'aire de Broca ont pu être observées lors de la lecture labiale. Ceci est cohérent avec de précédents résultats qui ont démontré une implication de l'aire de Broca pendant la discrimination phonétique de la parole, non seulement en condition auditive (Démonet *et al.*, 1992; Ojanen *et al.*, 2005; Vigneau *et al.*, 2006) mais également en condition visuelle (MacSweeney *et al.*, 2000; Campbell *et al.*, 2001; Paulesu *et al.*, 2003; Sadato *et al.*, 2004; Möttönen *et al.*, 2005). Alors que le langage des signes peut déclencher une activation de l'aire de Broca chez des sujets sourds profonds (Neville *et al.*, 1998; Petitto *et al.*, 2000) il a pu être montré que, chez les patients sourds implantés cochléaires, l'activation de l'aire de Broca dépend du degré de récupération auditive (Green *et al.*, 2005) et est négativement corrélée avec la durée de la surdité (Lee *et al.*, 2007). De tels résultats concordent avec nos données, qui révèlent une réactivation progressive de l'aire de Broca tandis que les patients récupèrent graduellement leurs capacités auditives, depuis le jour de mise en marche de l'implant jusqu'à plusieurs mois après l'implantation, lorsque les performances de compréhension de la parole sont stabilisées (voir discussion détaillée ci-après).

Enfin, on peut également remarquer un foyer d'activation situé dans le cortex frontal inférieur postérieur droit, au niveau de l'aire homologue de Broca. Pour la même tâche, ce foyer n'est plus observé dans le cas des sujets normo-entendants. Les rôles fonctionnels joués par l'aire homologue de Broca étant diversifiés, on pourrait proposer plusieurs interprétations possibles pour cette observation. Etant donné l'implication particulière du cortex frontal inférieur droit dans les processus de recherche en mémoire épisodique (Paulesu *et al.*, 1993) (voir *« Réseau frontal inférieur droit »*, § 4.5.4, p. 95), l'interprétation la plus plausible de cette activation serait un processus de confrontation entre les hypothèses articulatoires inférées à partir de l'observation des mouvements des lèvres et les schémas articulatoires accumulés en mémoire à long terme lors

de l'expérience quotidienne de la lecture labiale. Dans ce cadre, l'identification du mot prononcé à partir de la dynamique des mouvements des lèvres serait facilitée par la mémorisation préalable d'associations mot/dynamique au cours de situations perceptives similaires.

13.6.2 Réorganisation cérébrale du traitement audiovisuel de la parole

Conformément aux résultats de la littérature, la stimulation auditive permet une activation du réseau cortical auditif pour le patient implanté comme pour le sujet contrôle, avec des activations auditives localisées dans le gyrus temporal supérieur (Herzog *et al.*, 1991; Ito *et al.*, 1993; Truy *et al.*, 1995; Naito *et al.*, 1995; Okazawa *et al.*, 1996) (voir *« Réactivation des voies auditives centrales »*, § 9.4.1, p. 201).

De façon intéressante, on peut remarquer une suractivation du gyrus supramarginal droit chez les patients implantés, en comparaison avec les contrôles normo-entendants. Chez le sujet normo-entendant, cette région intervient probablement dans le stockage transitoire des représentations phonologiques internes de la parole (Hickok et Poeppel, 2004; Skipper *et al.*, 2005; Vigneau *et al.*, 2006) (voir *« Intégration corticale audio-visuo-motrice de la parole »*, § 4.4.3, p. 86). Aussi cette observation pourrait indiquer la nécessité chez le patient implanté d'un traitement phonologique approfondi de la parole avant son interprétation linguistique.

On peut également observer une suractivation du gyrus fusiforme droit chez les patients implantés lors de la présentation de parole audiovisuelle. Il est possible que cette suractivation soit révélatrice d'un traitement visuel approfondi des informations de lecture labiale, afin de les intégrer correctement avec les informations auditives.

Il est intéressant de remarquer que chez les patients implantés, la présentation de parole visuelle induit une suractivation de l'aire de la voix, qui est une aire spécifiquement auditive, tandis que la présentation de parole audiovisuelle, lorsqu'elle est contrastée avec une tâche visuelle de référence, induit une suractivation de l'aire fusiforme des visages, qui est une aire spécifiquement visuelle. Nous avons donc mis en évidence un renforcement des activations intermodales après l'implantation cochléaire, qui pourrait faire partie des corrélats neuronaux expliquant l'optimalité de l'intégration audiovisuelle de la parole réalisée par les patients implantés.

D'autre part l'aire de la voix activée lors de la présentation de parole visuelle est située dans l'hémisphère droit, de la même façon que l'aire fusiforme des visages activée lors de la présentation de parole audiovisuelle contrastée avec une tâche visuelle de référence est latéralisée à droite. Par conséquent les activations intermodales observées chez les patients implantés semblent plus présentes dans l'hémisphère droit, ce qui soulève une problématique intéressante qui sera examinée dans la suite.

Enfin, on observe lors de la stimulation par de la parole audiovisuelle une suractivation du cortex orbitofrontal chez les patients implantés (région frontale médiale ventrale). D'après Allison *et al.* (2000), cette région jouerait un rôle important dans la cognition et la perception sociale. En particulier, des activations de cette région par la vision de visages ont été démontrées par des études en enregistrements unitaires chez le singe (Thorpe *et al.*, 1983; Scalaidhe *et al.*, 1999), ainsi que par des études chez l'homme en potentiels évoqués (Allison *et al.*, 1999) et en tomographie par émission de positrons (Wicker *et al.*, 1998). La région orbitofrontale serait ainsi impliquée dans un réseau cérébral de cognition sociale faisant intervenir le sillon temporal supérieur, le cortex orbitofrontal et l'amygdale, toutes ces régions étant reliées entre elles par des connexions réciproques (Allison *et al.*, 2000). Ainsi, au sein de ce réseau, le cortex orbitofrontal jouerait un rôle dans le renforcement social (Rolls, 2000).

13.6.3 Rôle de l'aire temporale de la voix chez le patient implanté

Activations supranormales dans l'aire de la voix

Les patients implantés cochléaires montrent une activation significative de la partie la plus antérieure du gyrus temporal supérieur droit pour la plupart des contrastes visuels et audiovisuels, tandis que les contrôles normo-entendants ne présentent pas d'activation significative de cette région. Lorsque l'on compare les activations des patients et des contrôles, cette région apparaît comme significativement suractivée chez les patients pour les contrastes visuels et audiovisuels avec la condition de repos (V-R et AV-R), avec une tendance à la diminution de cette activité en fonction du temps d'utilisation de l'implant. Cette partie du cortex est une région auditive associative faisant partie du réseau des aires temporales de la voix, connu pour répondre spécifiquement à la voix humaine (Belin et Zatorre, 2003; Fecteau *et al.*, 2004a,b; von Kriegstein et Giraud, 2004) (voir *« Reconnaissance de la voix humaine »*, § 4.2.1, p. 81). En particulier, la partie la plus antérieure de ce réseau est spécifiquement impliquée lors de la reconnaissance de l'identité de la voix du locuteur.

Chez les patients implantés cochléaires, on assiste donc à des phénomènes de réorganisation corticale impliquant l'aire temporale de la voix, cette aire étant significativement suractivée lors d'une stimulation par de la parole visuelle ou audiovisuelle (par comparaison avec les sujets normo-entendants). Il est intéressant de noter que les capacités de discrimination des voix humaines sont sévèrement atteintes chez les patients implantés (Fu *et al.*, 2004, 2005), en particulier lorsque ces patients sont testés peu de temps après la mise en marche de l'implant. Par ailleurs, une étude récente en TEP a démontré que chez les patients implantés n'ayant pas encore récupéré leurs capacités de compréhension de la parole, des stimuli de voix humaine ne déclenchent aucune activation dans les aires temporales de la voix (Coez *et al.*, 2008). Cet ensemble de données suggère que, lors d'une période prolongée de surdité, la déprivation auditive rend inopérant le réseau cortical de traitement de la voix, celui-ci devenant alors susceptible de subir une réorganisation intermodale qui pourrait se traduire par une participation au traitement visuel de la parole. Cette hypothèse sera discutée en détail dans la suite de la discussion.

Participation de l'aire de la voix au traitement visuel de la parole

Nos données montrent clairement une suractivation significative de la partie antérieure du réseau des aires temporales de la voix lors d'une stimulation par de la lecture labiale chez les patients implantés inexpérimentés (T0). L'activation constatée est donc de nature intermodale, les aires temporales de la voix étant jusqu'ici connues pour ne répondre qu'à des stimuli auditifs (Belin *et al.*, 2000) et non pas à des stimuli visuels comme un visage humain (Thompson *et al.*, 2007). Ce résultat est à rapprocher d'autres études ayant démontré un couplage fonctionnel entre l'aire temporale de la voix et l'aire fusiforme des visages lors d'associations voix/visages (von Kriegstein *et al.*, 2005, 2006), ce qui fournirait une possible explication à l'activation intermodale de l'aire de la voix observée chez les patients implantés cochléaires.

Afin d'approfondir cette hypothèse, nous avons inclus l'aire temporale de la voix et l'aire fusiforme des visages dans notre analyse de connectivité fonctionnelle. Alors que le gyrus fusiforme présente des activations dans tous les groupes de sujets lors d'une stimulation par de la parole visuelle, aucune corrélation significative n'a pu être mise en évidence entre les niveaux d'activité observés dans ces deux aires. Cet absence de couplage fonctionnel rend improbable une implication de l'aire temporale de la voix dans un traitement non phonologique du visage humain (comme celui qui est effectué dans l'aire fusiforme des visages), et suggère que cette aire pourrait intervenir à des niveaux de traitement supérieurs, spécifiques de la parole, chez les

patients implantés inexpérimentés. Cette idée est soutenue par nos données, qui démontrent une nette corrélation positive à T0 entre le débit sanguin cérébral dans l'aire de la voix et les scores de reconnaissance obtenus par les patients en lecture labiale, que ces scores soient mesurés pendant la tomographie ou en séance orthophonique. Il est donc plausible que l'aire de la voix joue un rôle fonctionnel lors du traitement visuel de la parole chez les patients implantés cochléaires inexpérimentés, possiblement par l'intermédiaire d'interactions intermodales facilitatrices impliquant des stratégies d'association audiovisuelle.

Les données en neuroimagerie disponibles dans la littérature concordent avec cette hypothèse. Ainsi, l'étude des interactions intermodales chez les patients sourds postlinguaux implantés cochléaires démontre qu'une stimulation par de la parole auditive déclenche une activation intermodale des aires visuelles de bas niveau (BA 17/18) plus importante que chez les contrôles normo-entendants (Giraud *et al.*, 2001; Zatorre, 2001). Tandis que les rôles respectifs de l'audition et de la vision sont inversés dans notre étude, Giraud *et al.* (2001) interprètent leurs données d'une façon similaire, suggérant l'apprentissage de fortes associations audiovisuelles lors de la récupération de la parole après implantation cochléaire afin d'expliquer ces activations intermodales supranormales. Dans leur étude, de façon particulièrement intéressante, les performances en lecture labiale sont significativement corrélées avec les activations intermodales des aires visuelles par de la parole auditive, tandis que dans notre étude les performances en lecture labiale sont significativement corrélées avec les activations intermodales des aires auditives par de la parole visuelle. Ces données concordent pour indiquer, après implantation cochléaire, l'intervention d'activations intermodales supranormales des réseaux visuels par les réseaux auditifs (et vice-versa), les aires temporales de la voix jouant un rôle important dans le traitement de ces flux d'informations intermodales. Ce renforcement cortical des interactions intermodales serait alors directement impliqué dans l'amélioration progressive de la compréhension auditive (et dans une moindre mesure, visuelle) de la parole. Tout ceci tend à être confirmé par nos données comportementales, qui démontrent une intégration audiovisuelle supranormale de la parole chez les patients implantés cochléaires (*« Etudes comportementales chez le patient implanté »*, ch. 12, p. 257).

Dans ce contexte, les activations intermodales rapportées ici et dans les autres études pourraient traduire la reconstruction progressive d'une représentation multimodale de l'identité de l'autre après l'implantation cochléaire. Cette reconstruction ferait ainsi intervenir une association progressive des informations visuelles provenant du visage avec les informations auditives artificielles et dégradées provenant des sons de voix traités par l'implant cochléaire. Une telle reconstruction multimodale est un préliminaire nécessaire pour un traitement approprié de la parole audiovisuelle.

13.6.4 Rôle du réseau audio-visuo-moteur chez le patient implanté

Activations atypiques dans l'aire de Broca et la jonction pariéto-temporo-occipitale gauche

Les contrôles normo-entendants montrent une activation significative de l'aire de Broca (partie postérieure gauche des gyri frontaux inférieur et moyen) lors des stimulations par de la parole visuelle et audiovisuelle, tandis que les patients implantés ne présentent pas d'activation significative de cette région. Lorsque l'on compare les activations de groupe, l'aire de Broca apparaît comme significativement sous-activée chez les patients pour les contrastes visuels et audiovisuels avec la condition de repos (V-R et AV-R), avec une tendance à l'augmentation de l'activité en fonction du temps d'utilisation de l'implant.

D'autre part, chez les patients implantés cochléaires, la partie postérieure du sulcus temporal supérieur gauche (jonction pariéto-temporo-occipitale, ou PTO) était nettement impliquée pendant la tâche de lecture labiale à T0 aussi bien qu'à T1. Cette aire corticale joue un rôle crucial lors du traitement de la parole en participant à un encodage amodal de l'information phonologique, ce qui permet des traitements de plus haut-niveau comme un encodage sémantique (*sound-meaning interface*, Hickok et Poeppel (2004, 2007)). Cette région n'était au contraire pas impliquée lors de la lecture labiale chez les contrôles normo-entendants.

Ainsi, la suractivation observée dans la jonction PTO chez les patients implantés indiquerait que ces derniers sont capables d'utiliser des stratégies de traitement phonologique lors de la lecture labiale, ce qui leur permettrait d'atteindre de meilleures performances, tandis que les contrôles normo-entendants ne parviendraient que difficilement à associer les mouvements des lèvres avec les phonèmes qui leur correspondent. Ce raisonnement est cohérent avec les différences de performances en lecture labiale observées entre les patients et les contrôles. Les différences d'activation observées dans l'aire de Broca indiqueraient que les sujets contrôles associent de façon préférentielle les mouvements des lèvres avec une représentation motrice interne (activation de Broca, voir « *Intégration corticale audio-visuo-motrice de la parole* », § 4.4.3, p. 86) plutôt qu'avec une représentation phonologique interne (pas d'activation de la jonction PTO). Le phénomène inverse se produirait alors chez les patients implantés, la priorité étant donné à l'encodage phonologique des percepts audiovisuels, plutôt qu'à leur encodage moteur.

De façon intéressante, nous avons trouvé des corrélations positives significatives entre les niveaux d'activité observés dans la partie postérieure du sulcus temporal supérieur gauche et l'aire de Broca lors de la lecture labiale. Ces corrélations sont présentes dans les groupes de patients implantés cochléaires à T0 et T1, mais pas dans le groupe de contrôles normo-entendants. Cela indique que les patients implantés cochléaires présentent une coopération audio-visuo-motrice entre les représentations phonologiques induites à partir de l'information audiovisuelle (jonction PTO gauche) et les représentations motrices articulatoires correspondantes (aire de Broca) qui potentialise les performances de lecture labiale. Au contraire, chez les sujets normo-entendants, cette boucle audio-visuo-motrice est très peu valorisée lors de la lecture labiale, la jonction PTO gauche n'étant caractérisée par aucune activation significative ni aucune corrélation fonctionnelle avec l'aire de Broca.

Réactivation progressive du réseau audio-visuo-moteur de traitement de la parole

Plusieurs études ayant suggéré une corrélation entre le degré de métabolisme observé dans les aires normalement impliquées dans le traitement spécifique du langage, en particulier l'aire de Broca, et le degré de récupération perceptive chez les patients implantés (Okazawa *et al.*, 1996; Fujiki *et al.*, 1999), on peut envisager que les sous-activations observées tendent à disparaître à long terme, à mesure que les patients implantés récupèrent leurs capacités de compréhension de la parole. On peut observer une tendance en ce sens en condition de parole visuelle (V-RM) à T1 (voir fig. 13.10, p. 301), avec un foyer significativement suractivé dans l'aire de Broca ($p = 0.033$), tandis qu'aucune activation n'apparaît comme significative dans cette région à T0 ; toutefois la différence entre T0 et T1 n'apparaît pas comme statistiquement significative.

Ainsi, l'augmentation progressive du niveau d'activité dans l'aire de Broca pendant la lecture labiale chez les patients implantés suggère une réactivation du réseau audio-visuo-moteur de traitement de parole à mesure que les patients acquièrent de l'expérience dans l'utilisation de leur implant. Cette boucle liant perception et production de la parole implique des connexions réciproques entre les aires temporales supérieures postérieures et les régions préfrontales inférieures par l'intermédiaire du faisceau arqué (Binder et Frost, 1998; Belin *et al.*, 2000; Hickok et

Poeppel, 2000), permettant un échange d'information entre les réseaux temporaux postérieurs de traitement phonologique et les réseaux frontaux inférieurs de traitement articulatoire. Pendant la période de déprivation auditive, la correspondance entre la perception et la production de la parole est détériorée par la perte des informations auditives, si bien que l'interaction perception/production, qui implique les réseaux de neurones miroirs situés dans les régions préfrontales (Nishitani et Hari, 2002; Möttönen *et al.*, 2005; Skipper *et al.*, 2007) subissent une désactivation progressive pendant la surdité. Par ailleurs, plus les régions préfrontales présentent un hypermétabolisme au repos avant l'implantation cochléaire, plus la récupération de la compréhension de la parole est efficace après l'implantation cochléaire (Oh *et al.*, 2003; Kang *et al.*, 2004; Lee *et al.*, 2005, 2007).

Tout ceci montre l'importance du réseau audio-visuo-moteur pour la récupération de la perception de la parole après implantation cochléaire, dans la mesure où ce réseau de perception/production permet d'associer étroitement les représentations phonologiques avec les commandes articulatoires susceptibles de générer les phonèmes correspondants, dans le but d'aboutir à une représentation interne motrice non-ambiguë de la parole (Liberman *et al.*, 1967; Liberman et Mattingly, 1985; Liberman et Whalen, 2000). De la même façon que la boucle audio-visuo-motrice est susceptible de jouer un rôle clé pendant l'enfance, lors du développement de la parole (Hickok et Poeppel, 2000), il est probable qu'elle intervient de façon similaire lors de la réacquisition de la perception de la parole après une longue période de surdité.

13.6.5 Evolution post-implantation des stratégies de traitement de la parole

Différents niveaux de complexité de traitement de la parole

De précédentes études chez les sourds postlinguaux implantés cochléaires ont montré que le degré d'hypométabolisme temporal au repos avant implantation est prédictif de la récupération de la perception de la parole après implantation (Lee *et al.*, 2001). D'autres études menées chez les enfants sourds prélinguaux ont montré que le degré de récupération de la perception de la parole après implantation est corrélé au métabolisme pré-implantation, en particulier à l'hypermétabolisme dans les régions fronto-pariétales (comprenant l'aire de Broca) et l'hypométabolisme dans les régions fronto-pariétales (Oh *et al.*, 2003; Kang *et al.*, 2004; Lee *et al.*, 2005, 2007). Ces études démontrent que la récupération de la parole post-implantation est facilitée chez les sujets utilisant en priorité des réseaux de traitement de haut niveau de la parole (exécutifs, attentionnels, prédictifs et intégratifs) se situant dans les régions préfrontales et pariétales. Elles montrent également qu'au contraire, les sujets utilisant en priorité des réseaux de traitement de bas niveau, plus spécialisés dans l'analyse perceptive audiovisuelle de la parole (comme ceux présents dans les régions temporo-occipitales) rencontrent de plus grandes difficultés dans la récupération de la perception de la parole.

Cela indique que l'analyse perceptive audiovisuelle de la parole est nécessaire mais pas suffisante pour fournir une récupération optimale de la parole, et qu'elle doit être complétée par des analyses du plus haut niveau impliquant une focalisation de l'attention et un usage de la mémoire à court terme. Par conséquent, les réorganisations corticales qui peuvent être observées chez les patients implantés cochléaires lors de la lecture labiale (la modalité visuelle étant la moins évidente pour percevoir la parole) pourraient refléter un basculement d'une stratégie de traitement audiovisuel de bas niveau de la parole (activations de l'aire de la voix à T0 en lecture labiale) vers une stratégie mixte mêlant ce traitement audiovisuel avec des traitements attentionnels de haut niveau, qui permettent de mettre à profit des indices contextuels lexicaux et sémantiques afin d'atteindre un bon niveau de compréhension de la parole.

Latéralisation du traitement de la parole

De façon générale, toutes conditions confondues, on peut remarquer que les régions significativement suractivées chez les contrôles normo-entendants sont regroupées dans les réseaux auditifs, associatifs et cognitifs de l'hémisphère gauche, tandis que les régions suractivées chez les patients implantés sont regroupées dans les réseaux associatifs et visuels de l'hémisphère droit. En particulier, lors d'une stimulation par de la parole audiovisuelle, on observe de façon remarquable chez les patients implantés un équilibrage des activations entre les deux hémisphères, avec une tendance en faveur de l'hémisphère droit. Au contraire, dans le cas des sujets normo-entendants, les activations auditives sont nettement latéralisées dans l'hémisphère gauche. Enfin, comme cela a été souligné précédemment, les activations intermodales observées chez les patients implantés se limitent à l'hémisphère droit.

On peut interpréter ces observations de deux façons différentes. D'une part, l'hémisphère gauche étant spécialisé dans le traitement du langage, tout particulièrement chez les sujets droitiers (Démonet *et al.*, 1992; Klein *et al.*, 1995; Zatorre *et al.*, 1996; Frost *et al.*, 1999; Rutten *et al.*, 2002; Hirata *et al.*, 2004; Kirveskari *et al.*, 2006) (voir *« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91), cette différence hémisphérique peut être révélatrice d'un traitement supplémentaire de nature non linguistique réalisé chez les patients implantés lors de la compréhension de la parole visuelle ou audiovisuelle. Dans le cas de la parole visuelle, cela peut indiquer la nécessité d'un traitement cortical plus approfondi afin de faire correspondre les mouvements visuels des lèvres et du visage avec des sons de la voix délivrés par l'implant cochléaire et placés en mémoire à long terme.

D'autre part, l'hémisphère droit serait spécialisé dans l'analyse spectrale fine des signaux sonores (Zatorre *et al.*, 2002) (voir également *« Latéralisation hémisphérique du langage »*, § 4.5.1, p. 91). Comme le signal délivré par l'implant cochléaire est riche en informations temporelles mais relativement pauvre en informations fréquentielles (voir *« Rôle de l'enveloppe temporelle et de la structure fine »*, § 9.2.2, p. 194), il nécessite un traitement spectral approfondi qui est probablement réalisé dans l'hémisphère droit chez les patients implantés, ce qui expliquerait la sous-activation de l'hémisphère gauche et la suractivation de l'hémisphère droit observées chez ces derniers.

Il est intéressant de remarquer que les suractivations observées chez les implantés lors de leur première tomographie (T0) sont fortement latéralisées dans l'hémisphère droit (voir fig. 13.15, p. 305 et fig. 13.16, p. 305). Par ailleurs la comparaison des profils d'activation entre T0 et T1 (voir tab. 13.3, p. 307) laisse soupçonner une évolution à long terme de la latéralisation du traitement de la parole chez l'implanté cochléaire. Ainsi, dans les premières semaines suivant la mise en marche de l'implant cochléaire, cette latéralisation serait déséquilibrée vers l'hémisphère droit, incluant en particulier une suractivation des aires auditives associatives droites, une suractivation du lobe temporal antérieur droit (dont l'aire de la voix) et probablement du lobe frontal inférieur droit. Ensuite, à long terme et une fois le processus de réorganisation corticale stabilisé, la latéralisation se rééquilibrerait en faveur de l'hémisphère gauche, incluant une réactivation de l'aire de Broca ainsi qu'un équilibrage bilatéral des activations dans les aires auditives associatives. Ces observations sont cohérentes avec l'hypothèse d'une implication de cet hémisphère dans un traitement spectral approfondi des sons dégradés délivrés par l'implant cochléaire. En effet l'optimisation du traitement effectué par ces réseaux neuronaux implique de plus fortes activations de ces réseaux au début de la période d'utilisation de l'implant, afin d'assurer une réorganisation des connexions synaptiques adaptée à cette nouvelle forme de traitement. Au contraire, après une longue période d'utilisation de l'implant, les réseaux neuronaux sont déjà adaptés par réorganisation corticale et consomment moins d'énergie pour traiter l'information.

Bibliographie

Allison, T., Puce, A. et McCarthy, G. (2000). Social perception from visual cues : role of the STS region. *Trends Cogn. Sci.*, 4(7):267–278.

Allison, T., Puce, A., Spencer, D. D. et McCarthy, G. (1999). Electrophysiological studies of human face perception. I : Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex*, 9(5):415–430.

Belin, P. et Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, 14(16):2105–2109.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P. et Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403(6767):309–312.

Bernstein, L. E., Auer, Jr., E. T., Moore, J. K., Ponton, C. W., Don, M. et Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, 13:311–315.

Binder, J. R. et Frost, J. A. (1998). Functional MRI studies of language processes in the brain. *Neurosci. News*, 1:15–23.

Campbell, R. et Capek, C. (2008). Seeing speech and seeing sign : insights from a fMRI study. *Int. J. Audiol.*, 47 Suppl 2:S3–S9.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J. et David, A. S. (2001). Cortical substrates for the perception of face actions : an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Brain Res. Cogn. Brain Res.*, 12(2):233–243.

Capek, C. M., Macsweeney, M., Woll, B., Waters, D., McGuire, P. K., David, A. S., Brammer, M. J. et Campbell, R. (2008). Cortical circuits for silent speechreading in deaf and hearing people. *Neuropsychologia*, 46(5):1233–1241.

Coez, A., Zilbovicius, M., Ferrary, E., Bouccara, D., Mosnier, I., Ambert-Dahan, E., Bizaguet, E., Syrota, A., Samson, Y. et Sterkers, O. (2008). Cochlear implant benefits in deafness rehabilitation : PET study of temporal voice activations. *J. Nucl. Med.*, 49(1):60–67.

Content, A., Mousty, P. et Radeau, M. (1990). Brulex. une base de données lexicales informatisée pour le français écrit et parlé. *Ann. Psychol.*, 90:551–566.

Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A. et Frackowiak, R. S. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115(Pt 6):1753–1768.

Démonet, J.-F., Thierry, G. et Cardebat, D. (2005). Renewal of the neurophysiology of language : functional neuroimaging. *Physiol. Rev.*, 85(1):49–95.

Fecteau, S., Armony, J. L., Joanette, Y. et Belin, P. (2004a). Is voice processing species-specific in human auditory cortex? an fMRI study. *Neuroimage*, 23(3):840–848.

Fecteau, S., Armony, J. L., Joanette, Y. et Belin, P. (2004b). Priming of non-speech vocalizations in male adults : the influence of the speaker's gender. *Brain Cogn.*, 55(2):300–302.

Frost, J. A., Binder, J. R., Springer, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M. et Cox, R. W. (1999). Language processing is strongly left lateralized in both sexes. Evidence from functional MRI. *Brain*, 122(Pt 2):199–208.

Fu, K. M., Shah, A. S., O'Connell, M. N., McGinnis, T., Eckholdt, H., Lakatos, P., Smiley, J. et Schroeder, C. E. (2004). Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. *J. Neurophysiol.*, 92(6):3522–3531.

Fu, Q. J., Nogaki, G. et Galvin, 3rd, J. J. (2005). Auditory training with spectrally shifted speech : implications for cochlear implant patient auditory rehabilitation. *J. Assoc. Res. Otolaryngol.*, 6(2):180–189.

Fujiki, N., Naito, Y., Hirano, S., Kojima, H., Shiomi, Y., Nishizawa, S., Konishi, J. et Honjo, I. (1999). Correlation between rCBF and speech perception in cochlear implant users. *Auris Nasus Larynx*, 26(3):229–236.

Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E. et Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, 30(3):657–663.

Green, K. M., Julyan, P. J., Hastings, D. L. et Ramsden, R. T. (2005). Auditory cortical activation and speech perception in cochlear implant users : effects of implant experience and duration of deafness. *Hear. Res.*, 205(1-2):184–192.

Haxby, J. V., Hoffman, E. A. et Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.*, 4(6):223–233.

Herzog, H., Lamprecht, A., Kuhn, A., Roden, W., Vosteen, K. H. et Feinendegen, L. E. (1991). Cortical activation in profoundly deaf patients during cochlear implant stimulation demonstrated by H_2O^{15} PET. *J. Comput. Assist. Tomogr.*, 15(3):369–375.

Hickok, G. et Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.*, 4(4):131–138.

Hickok, G. et Poeppel, D. (2004). Dorsal and ventral streams : a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99.

Hickok, G. et Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.*, 8(5):393–402.

Hirata, M., Kato, A., Taniguchi, M., Saitoh, Y., Ninomiya, H., Ihara, A., Kishima, H., Oshino, S., Baba, T., Yorifuji, S. et Yoshimine, T. (2004). Determination of language dominance with synthetic aperture magnetometry : comparison with the Wada test. *Neuroimage*, 23(1):46–53.

Ito, J., Sakakibara, J., Iwasaki, Y. et Yonekura, Y. (1993). Positron emission tomography of auditory sensation in deaf patients and patients with cochlear implants. *Ann. Otol. Rhinol. Laryngol.*, 102(10):797–801.

Kang, E., Lee, D. S., Kang, H., Lee, J. S., Oh, S. H., Lee, M. C. et Kim, C. S. (2004). Neural changes associated with speech learning in deaf children following cochlear implantation. *Neuroimage*, 22(3):1173–1181.

Kanwisher, N., McDermott, J. et Chun, M. M. (1997). The fusiform face area : a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, 17(11):4302–4311.

Kanwisher, N. et Yovel, G. (2006). The fusiform face area : a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 361(1476):2109–2128.

Kirveskari, E., Salmelin, R. et Hari, R. (2006). Neuromagnetic responses to vowels vs. tones reveal hemispheric lateralization. *Clin. Neurophysiol.*, 117(3):643–648.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E. et Evans, A. C. (1995). The neural substrates underlying word generation : a bilingual functional-imaging study. *Proc. Natl. Acad. Sci. U.S.A.*, 92(7):2899–2903.

Lambertz, N., Gizewski, E. R., de Greiff, A. et Forsting, M. (2005). Cross-modal plasticity in deaf subjects dependent on the extent of hearing loss. *Brain Res. Cogn. Brain Res.*, 25(3):884–890.

Lee, D. S., Lee, J. S., Oh, S. H., Kim, S. K., Kim, J. W., Chung, J. K., Lee, M. C. et Kim, C. S. (2001). Cross-modal plasticity and cochlear implants. *Nature*, 409(6817):149–150.

Lee, H. J., Giraud, A.-L., Kang, E., Oh, S. H., Kang, H., Kim, C. S. et Lee, D. S. (2007). Cortical activity at rest predicts cochlear implantation outcome. *Cereb. Cortex*, 17(4):909–917.

Lee, H. J., Kang, E., Oh, S. H., Kang, H., Lee, D. S., Lee, M. C. et Kim, C. S. (2005). Preoperative differences of cerebral metabolism relate to the outcome of cochlear implants in congenitally deaf children. *Hear. Res.*, 203:2–9.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. et Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.*, 74(6):431–461.

Liberman, A. M. et Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1):1–36.

Liberman, A. M. et Whalen, D. H. (2000). On the relation of speech to language. *Trends Cogn. Sci.*, 4(5):187–196.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., Williams, S. C., Woll, B. et Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise : An event-related fMRI study. *Neuroreport*, 11:1729–1733.

McCarthy, G., Puce, A., Gore, J. C. et Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *J. Cogn. Neurosci.*, 9:605–610.

Möttönen, R., Jarvelainen, J., Sams, M. et Hari, R. (2005). Viewing speech modulates activity in the left SI mouth cortex. *Neuroimage*, 24(3):731–737.

Naito, Y., Okazawa, H., Honjo, I., Hirano, S., Takahashi, H., Shiomi, Y., Hoji, W., Kawano, M., Ishizu, K. et Yonekura, Y. (1995). Cortical activation with sound stimulation in cochlear implant users demonstrated by positron emission tomography. *Cogn. Brain Res.*, 2(3):207–214.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P. et Turner, R. (1998). Cerebral organization for language in deaf and hearing subjects : biological constraints and effects of experience. *Proc. Natl. Acad. Sci. U.S.A.*, 95(3):922–929.

Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. et Kubo, T. (1999). Sign language 'heard' in the auditory cortex. *Nature*, 397(6715):116.

Nishitani, N. et Hari, R. (2002). Viewing lip forms : cortical dynamics. *Neuron*, 36(6):1211–1220.

Oh, S. H., Kim, C. S., Kang, E. J., Lee, D. S., Lee, H. J., Chang, S. O., Ahn, S. H., Hwang, C. H., Park, H. J. et Koo, J. W. (2003). Speech perception after cochlear implantation over a 4-year time period. *Acta Otolaryngol.*, 123(2):148–153.

Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I. P., Joensuu, R., Autti, T. et Sams, M. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage*, 25(2):333–338.

Okazawa, H., Naito, Y., Yonekura, Y., Sadato, N., Hirano, S., Nishizawa, S., Magata, Y., Ishizu, K., Tamaki, N., Honjo, I. et Konishi, J. (1996). Cochlear implant efficiency in pre- and postlingually deaf subjects. A study with H_2O^{15} and PET. *Brain*, 119:4.

Paulesu, E., Frith, C. D. et Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362(6418):342–345.

Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., Sensolo, S. et Fazio, F. (2003). A functional-anatomical model for lipreading. *J. Neurophysiol.*, 90(3):2005–2013.

Petitto, L. A. et Kovelman, I. (2003). The Bilingual Paradox : How signing-speaking bilingual children help us to resolve it and teach us about the brain's mechanisms underlying all language acquisition. *Learning. Languages.*, 8(3):5–18.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D. et Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages : implications for the neural basis of human language. *Proc. Natl. Acad. Sci. U.S.A.*, 97(25):13961–13966.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cereb. Cortex*, 10(3):284–294.

Rutten, G. J., Ramsey, N. F., van Rijen, P. C., Alpherts, W. C. et van Veelen, C. W. (2002). fMRI-determined language lateralization in patients with unilateral or mixed language dominance according to the Wada test. *Neuroimage*, 17(1):447–460.

Sadato, N., Yamada, H., Okada, T., Yoshida, M., Hasegawa, T., Matsuki, K., Yonekura, Y. et Itoh, H. (2004). Age-dependent plasticity in the superior temporal sulcus in deaf humans : a functional MRI study. *BMC Neurosci.*, 5(1):56.

Scalaidhe, S. P., Wilson, F. A. et Goldman-Rakic, P. S. (1999). Face-selective neurons during passive viewing and working memory performance of rhesus monkeys : evidence for intrinsic specialization of neuronal coding. *Cereb. Cortex*, 9(5):459–475.

Skipper, J. I., Nusbaum, H. C. et Small, S. L. (2005). Listening to talking faces : motor cortical activation during speech perception. *Neuroimage*, 25(1):76–89.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C. et Small, S. L. (2007). Hearing lips and seeing voices : how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex*.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D. et Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage*, 37(3):966–973.

Thorpe, S. J., Rolls, E. T. et Maddison, S. (1983). The orbitofrontal cortex : neuronal activity in the behaving monkey. *Exp. Brain Res.*, 49(1):93–115.

Truy, E., Deiber, M. P., Cinotti, L., Mauguiere, F., Froment, J. C. et Morgon, A. (1995). Auditory cortex activity changes in long-term sensorineural deprivation during crude cochlear electrical stimulation : evaluation by positron emission tomography. *Hear. Res.*, 86(1-2):34–42.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B. et Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas : phonology, semantics, and sentence processing. *Neuroimage*, 30(4):1414–1432.

von Kriegstein, K. et Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2):948–955.

von Kriegstein, K., Kleinschmidt, A. et Giraud, A.-L. (2006). Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cereb. Cortex*, 16(9):1314–1322.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P. et Giraud, A.-L. (2005). Interaction of face and voice areas during speaker recognition. *J. Cogn. Neurosci.*, 17(3):367–376.

Wicker, B., Michel, F., Henaff, M. A. et Decety, J. (1998). Brain regions involved in the perception of gaze : a PET study. *Neuroimage*, 8(2):221–227.

Zatorre, R. J. (2001). Do you see what I'm saying ? interactions between auditory and visual cortices in cochlear implant users. *Neuron*, 31(1):13–14.

Zatorre, R. J., Belin, P. et Penhune, V. B. (2002). Structure and function of auditory cortex : music and speech. *Trends Cogn. Sci.*, 6(1):37–46.

Zatorre, R. J., Meyer, E., Gjedde, A. et Evans, A. C. (1996). PET studies of phonetic processing of speech : review, replication, and reanalysis. *Cereb. Cortex*, 6(1):21–30.

Conclusion générale

Aspects dynamiques de la réorganisation neurofonctionnelle

Nos données comportementales, en appui avec les observations rapportées dans la littérature scientifique, confirment l'existence d'une évolution rapide des performances auditives et audiovisuelles dans les premiers mois suivant l'implantation cochléaire. Nous observons en particulier une stabilisation très rapide de la perception audiovisuelle dès le deuxième mois après la mise en marche de l'implant, tandis que la stabilisation de la perception auditive se fait à une vitesse plus modérée, à partir de sept mois après la première stimulation par l'implant.

Ces résultats montrent également que les compréhensions auditive et audiovisuelle de la parole se rapprochent de la normale à mesure de l'utilisation de l'implant, les patients implantés obtenant des scores audiovisuels supérieurs à 95% dès le troisième mois post-implantation. Toutefois les performances auditives à long terme se maintiennent autour d'un score de reconnaissance de 80%, ce qui indique que les informations auditives isolées ne permettent pas d'aboutir à une compréhension parfaite de la parole.

Pour finir, nous pouvons observer chez les patients implantés un maintien à long terme de compétences supra-normales en lecture labiale, avec un score moyen de 35%, plus de trois fois supérieur au score de 10% obtenu par les sujets normo-entendants. Nous démontrons que le maintien de ces performances visuelles supra-normales permet de compenser la dégradation spectrale des informations auditives traitées par l'implant cochléaire lors de la perception audiovisuelle de la parole (voir ci-après, *« Coopération intermodale entre vision et audition »*).

Les activations cérébrales obtenues dans notre protocole d'étude en imagerie fonctionnelle viennent renforcer ces observations comportementales. La réorganisation des réseaux neuronaux impliqués dans la perception audiovisuelle de la parole chez les patients implantés cochléaires présente des aspects dynamiques particulièrement marqués au niveau de la latéralisation hémisphérique, avec un surinvestissement de l'hémisphère droit dans le traitement audiovisuel de la parole, ce surcroît d'activation tendant ensuite à s'amenuiser à long terme. Nous avons vu précédemment que cette surimplication de l'hémisphère cérébral droit pourrait correspondre à la nécessité d'un traitement spectral approfondi du signal dégradé délivré par l'implant cochléaire.

Ainsi, les patients implantés testés peu après l'entrée en fonction de l'implant présentent une répartition inhabituelle des activations entre les deux hémisphères, avec une activation supranormale de l'hémisphère droit (dont en particulier le lobe temporal antérieur droit, incluant l'aire sélective à l'identité de la voix) accompagnée par une activation subnormale de l'hémisphère gauche (dont en particulier le lobe frontal inférieur gauche, incluant l'aire de Broca). Ces tendances sont moins accentuées à long terme, avec une diminution significative des suractivations du lobe temporal antérieur droit et une probable augmentation des activations dans le lobe frontal inférieur gauche.

Par conséquent le profil d'activation cérébral observé chez les patients implantés tend à long terme vers un profil d'activation plus caractéristique des normo-entendants,

sans toutefois la forte latéralisation dans l'hémisphère gauche que l'on peut classiquement observer chez ces derniers. Cette convergence du profil d'activation corticale vers un profil plus proche de la normale est à mettre en parallèle avec l'importante récupération perceptive mise en évidence au travers de nos études comportementales.

De façon générale, cet ensemble d'observations en comportement et en neuroimagerie cérébrale indique une forte mobilisation des processus de réorganisation neurofonctionnelle dans les premiers mois suivant la mise en marche de l'implant cochléaire, suivie par une stabilisation à long terme à la fin de la première année post-implantation. Ces données permettent de préciser la fenêtre temporelle critique pendant laquelle la rééducation orthophonique des patients implantés doit atteindre une efficacité optimale.

Coopération intermodale entre vision et audition

Lorsque nous comparons les performances de compréhension audiovisuelle de la parole des implantés avec des sujets normo-entendants placés dans des conditions de dégradation auditive similaires (simulation acoustique d'implant cochléaire), nous démontrons que les patients implantés sont capables d'intégrer de façon optimale les indices visuels et auditifs, tandis que l'intégration audiovisuelle réalisée par les sujets normo-entendants reste en-dessous du seuil d'optimalité. Pour un niveau lexical de traitement de la parole audiovisuelle, les patients implantés présentent donc un net avantage perceptif par rapport aux sujets normo-entendants.

Par ailleurs, l'utilisation de stimuli de parole audiovisuelle incongruente de type McGurk montre que pour des bas niveaux de traitement, comme le niveau phonétique, les mécanismes d'intégration audiovisuelle phonétique présentent certaines similarités de structure entre les patients implantés et les sujets normo-entendants. En effet, dans les deux cas, l'intégration audiovisuelle de la parole aboutit à la sélection du percept le mieux compatible du point de vue articulatoire avec les informations délivrées par l'audition et la vision. Par ailleurs les patients implantés ne montrent aucun avantage perceptif par rapport aux sujets normo-entendants lors de l'intégration audiovisuelle de la parole à un niveau phonétique.

Ainsi l'intégration audiovisuelle de la parole bénéficierait probablement chez les patients implantés d'influences provenant des hauts niveaux de traitement cognitif du langage (lexical, sémantique et syntaxique) aboutissant à un meilleur couplage audiovisuel, ce qui permettrait de compenser la perte d'informations auditives due au traitement spectral grossier réalisé par le processeur de l'implant cochléaire.

Ce renforcement de la coopération intermodale chez les patients implantés cochléaires se trouve confirmé par les observations effectuées en neuroimagerie cérébrale. En effet, la présentation de parole visuelle induit chez les patients implantés une suractivation des réseaux auditifs associatifs impliqués dans l'identification du locuteur (sillon temporal supérieur antérieur droit) et le traitement phonologique de la parole (sillon temporal supérieur postérieur bilatéral), tandis que ces suractivations ne sont pas observées dans le cas des contrôles normo-entendants. Par ailleurs la présentation de parole audiovisuelle induit chez les implantés une suractivation des réseaux visuels associatifs (gyrus fusiforme) que l'on ne retrouve pas dans le cas des sujets contrôles.

Le fort couplage audiovisuel observé chez les patients implantés est un premier fac-

teur permettant d'expliquer le haut niveau de performance audiovisuelle rapidement atteint par ces derniers. Un second facteur important est représenté par le surinvestissement de la modalité visuelle développé par les implantés cochléaires. Ce surinvestissement est démontré d'une part par la supranormalité des aptitudes en lecture labiale mise en évidence de façon très consistante au travers de l'évaluation comportementale des performances, et d'autre part par le déséquilibre de l'intégration audiovisuelle phonétique de la parole en faveur des informations visuelles, mis en évidence par l'utilisation de stimuli phonétiques audiovisuels incongruents (McGurk).

Dans ce dernier cas, l'observation de l'effet McGurk indique que l'intégration audiovisuelle de la parole chez le sujet sourd implanté cochléaire fait en réalité intervenir une survalorisation des indices articulatoires visuels et une sous-valorisation des indices articulatoires auditifs. Au contraire, les sujets normo-entendants attribuent des pertinences équilibrées aux indices auditifs et visuels. Cependant le mécanisme d'intégration globale reste le même, puisque l'intégration audiovisuelle se produit toujours par sélection des entités articulatoires les mieux compatibles avec les indices auditifs et visuels perçus. Seule la qualité des informations afférentes et les poids respectifs qui leur sont attribués sont modifiés chez les patients implantés.

Enfin, un dernier facteur est constitué par l'entraînement audiovisuel quotidien que subissent les patients en utilisant l'implant cochléaire pour converser avec leurs semblables. Le bénéfice d'un entraînement audiovisuel est suggéré fortement par l'avantage perceptif démontré lors de l'évaluation comportementale des patients, et démontré de façon consistante par notre étude réalisée chez des sujets normo-entendants soumis à un entraînement par une simulation acoustique d'implant cochléaire. Dans ce cas précis, les sujets entraînés en condition audiovisuelle présentent un avantage auditif par rapport aux sujets entraînés en condition auditive seule. Cette observation expérimentale démontre que le couplage audiovisuel en situation naturelle est un élément crucial non seulement dans la récupération audiovisuelle, mais également dans la récupération auditive seule.

De façon générale, nos données psychophysiques et neurofonctionnelles convergent pour indiquer l'extrême importance du couplage audiovisuel chez les patients implantés cochléaires, aussi bien au niveau de l'entraînement perceptif quotidien qu'à travers son rôle facilitateur lors de la perception audiovisuelle de la parole.

Différences implantation droite/gauche

L'existence hypothétique de différences entre les processus de réorganisation neurofonctionnelle engagés chez les patients implantés à droite ou à gauche est un sujet de débat important. À l'heure actuelle le côté à implanter est en effet sélectionné sur la base de critères comportementaux et otologiques, mais il n'existe aucun consensus général sur ces critères.

Nos données en neuroimagerie fonctionnelle fournissent peu d'éléments permettant de trancher la question. Elles montrent une suractivation du lobe frontal droit chez les patients implantés à gauche. Cette observation pourrait être interprétée dans le sens d'un avantage perceptif, étant donné la corrélation existant entre la suractivation du cortex préfrontal et la facilitation perceptive chez les patients implantés (« *Utilisation du métabolisme pré-implantation comme facteur prédictif* », § 9.3.4, p. 200). Par ailleurs, étant donné le rôle crucial joué par l'hémisphère droit dans les premiers mois d'utilisation

de l'implant cochléaire, il ne serait pas surprenant qu'une implantation à gauche favorise un traitement spectral approfondi par l'hémisphère droit, et donc joue un rôle facilitateur dans le processus de réorganisation neurofonctionnelle. Quoi qu'il en soit, l'apport de réponses pertinentes à cette problématique mérite des études plus approfondies.

Différences hommes/femmes

L'analyse approfondie des résultats obtenus par les patients implantés et les sujets contrôles normo-entendants en condition de lecture labiale fait apparaître de nettes différences perceptives entre hommes et femmes, les femmes étant systématiquement avantagées lors de la perception visuelle de la parole. Ces différences sont par ailleurs tout à fait cohérentes avec les données rapportées dans la littérature scientifique.

Cet avantage perceptif peut être expliqué d'une part par l'emploi de stratégies cognitives de haut niveau par les femmes, qui permettraient d'intégrer et de prédire les mots prononcés silencieusement par le locuteur, et d'autre part par une coopération audiovisuelle plus efficace chez les femmes, facilitée par des mécanismes de confrontation entre les indices visuels et les hypothèses auditives qui peuvent en découler.

Il est par ailleurs intéressant d'observer un amenuisement de cet avantage perceptif après implantation cochléaire, à mesure de l'utilisation de l'implant. Nous expliquons cette différence par le développement chez les hommes, lors de l'implantation cochléaire, de stratégies cognitives similaires à celles utilisées par les femmes dans une situation de normalité auditive, ce qui permet aux hommes de rattraper leur retard perceptif sur les femmes. Chez les hommes, le développement de ces stratégies spécifiques serait donc induit par la situation perceptive difficile que représente l'implantation cochléaire, ces stratégies n'étant pas privilégiées en situation de normalité auditive.

L'existence de différences neurofonctionnelles lors de la perception visuelle de la parole chez les hommes et les femmes nous conduit à proposer une différenciation des programmes de réhabilitation orthophonique après implantation cochléaire. Dans ce cadre les hommes nécessiteraient une réhabilitation favorisant tout particulièrement les exercices en lecture labiale afin de développer les stratégies cognitives compensatrices qu'ils ne privilégient pas de façon naturelle.

Annexes

Annexe A

Neuroanatomie

Sommaire

A.1 Description anatomique 339
A.1.1 Structures encéphaliques 339
A.1.2 Aires de Brodmann 341
A.2 Normalisation de l'anatomie corticale 347
A.2.1 Variabilité anatomique 347
A.2.2 Normalisation spatiale 348
A.2.3 Atlas de référence 348
Bibliographie 349

Introduction

Les diverses structures anatomiques qui constituent les hémisphères cérébraux peuvent être décrites de différentes manières. Une première façon de se repérer est d'utiliser les subdivisions naturelles du cortex en circonvolutions et en sillons (approche topographique). Une deuxième est de s'appuyer sur les similarités de l'organisation cellulaire des différentes couches du cortex (approche cytoarchitectonique par les aires de Brodmann). Enfin, une dernière méthode consiste à compartimenter le cerveau en régions fonctionnellement distinctes. Ces trois approches, concordantes sur certains points, ne sont toutefois pas équivalentes. L'approche neuroanatomique par régionalisation fonctionnelle ayant été en grande partie abordée dans les chapitres précédents, nous nous focaliserons essentiellement dans la suite sur les approches topographique et cytoarchitectonique.

A.1 Description anatomique

A.1.1 Structures encéphaliques[1]

Le cerveau[†], ou plus précisément l'encéphale, désigne la partie du système nerveux central logée dans la boîte crânienne. Le cerveau est enveloppé par une couche de tissu conjonctif, la pie-mère, qui tapisse la surface du cortex. La paroi interne de la voûte crânienne est quant à elle

[1] d'après Altenmuller (2003)

tapissée par une couche conjonctive épaisse, la dure-mère, elle-même recouverte par une deuxième couche conjonctive, l'arachnoïde. Cette dernière est séparée de la pie-mère par une couche épaisse de liquide cérébro-spinal (ou liquide céphalo-rachidien). Les trois enveloppes conjonctives, la pie-mère, l'arachnoïde et la dure-mère, appelées méninges, jouent avec le liquide cérébro-spinal et la voûte crânienne un rôle de protection physique, chimique et biologique pour l'encéphale.

Le cerveau humain pèse en moyenne 1200 grammes chez l'homme adulte. Il peut être subdivisé en trois parties : le rhombencéphale (*hindbrain*), le mésencéphale (*midbrain*) et le prosencéphale (*forebrain*). Le rhombencéphale est formé du bulbe rachidien (*medulla*), du pont (*pons*) et du cervelet† (*cerebellum*). Le bulbe rachidien, le pont et le mésencéphale forment ensemble le tronc cérébral† (*brainstem*), qui est la partie la plus primitive du cerveau du point de vue phylogénétique. Le tronc cérébral régule toutes les fonctions vitales comme la respiration, les battements cardiaques, la température du corps et l'équilibre. Il contrôle également de nombreuses fonctions sensorielles et motrices comme les mouvements des yeux et la coordination des réflexes visuels et auditifs. Le cervelet se trouve en arrière du pont et contrôle principalement l'équilibre du corps et la synchronisation des mouvements, mais peut également intervenir dans des fonctions cognitives complexes comme le langage (voir « *Traitement du langage dans le cervelet* », § 4.5.3, p. 94).

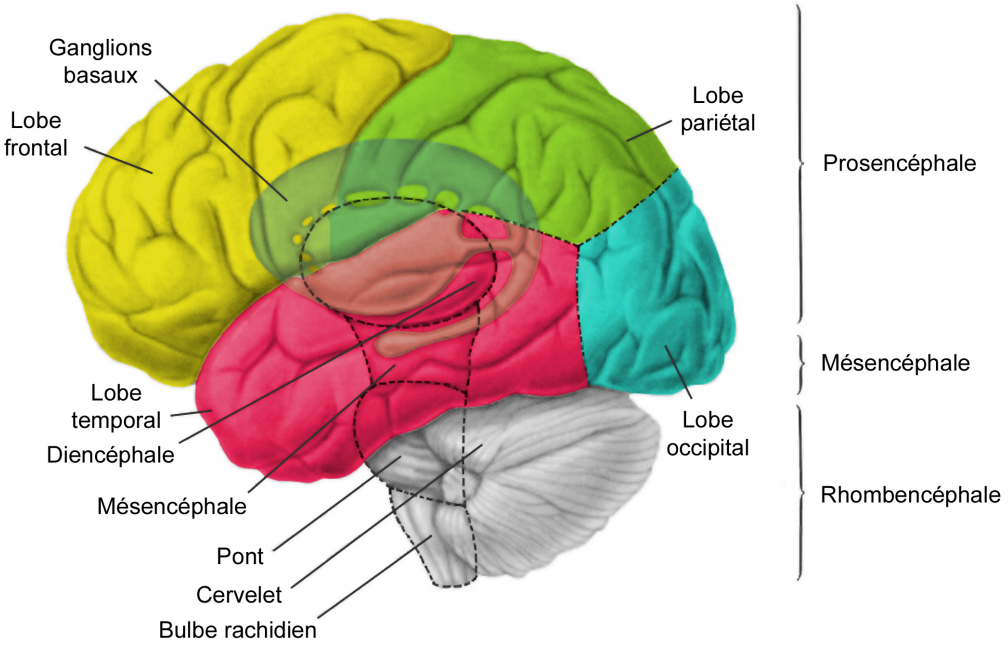


FIG. A.1 – Structures encéphaliques. Adapté d'après Altenmuller (2003).

Le prosencéphale est constitué du télencéphale, regroupant les deux hémisphères cérébraux, du diencéphale, regroupant le thalamus† et l'hypothalamus, et de trois structures situées en profondeur entre les hémisphères : les ganglions basaux, l'hippocampe et le noyau amygdaloïde (les deux dernières se trouvent à la frontière interne du lobe temporal et ne sont pas montrées sur la figure). Les ganglions basaux jouent un rôle dans la régulation des tâches motrices, l'hippocampe est associé à de nombreux aspects de la mémoire, tandis que le noyau amygdaloïde coordonne les reponses autonomes et endocrines avec les états émotionnels. Le thalamus permet de relayer les

informations afférentes en provenance de tous les systèmes sensoriels et les transmet au cortex cérébral, tandis que l'hypothalamus régule les fonctions autonomes et endocrines.

Toutes les fonctions cognitives sont régulées par le cortex cérébral†, la partie la plus extérieure du cerveau. Selon des estimations récentes, le cortex cérébral est constitué par approximativement cent milliards (10^{11}) de neurones reliés entre eux par un réseau dense de fibres nerveuses. Par l'intermédiaire des synapses, chaque cellule nerveuse peut communiquer avec approximativement dix mille (10^5) autres cellules nerveuses. Le cortex cérébral est divisé en deux hémisphères reliés entre eux par un regroupement de fibres contenant environ cent millions (10^8) de fibres nerveuses, appelé corps calleux (corps calleux). La surface du cerveau présente des circonvolutions (usuellement appelées *gyri*, singulier gyrus†) délimitées par des sillons (usuellement appelés *sulci*, singulier sulcus†, ou scissures ou fissures) de profondeur variable.

Quatre propriétés importantes caractérisent l'organisation du cortex cérébral. Tout d'abord, chaque hémisphère est principalement associé aux processus perceptifs et moteurs correspondant au côté contralatéral du corps. Ensuite, ces hémisphères, bien que similaires, ne sont ni complètement symétriques en structure ni équivalents du point de vue fonctionnel. D'autre part, le cortex est organisé hiérarchiquement, en aires motrices ou sensorielles primaires, secondaires et associatives (ou tertiaires) distinctes les unes des autres. Les aires sensorielles primaires sont reliées directement aux organes perceptifs par le biais du thalamus. Les aires motrices primaires sont reliées directement à la moelle épinière. Les aires secondaires et les aires associatives sont adjacentes aux aires primaires et effectuent des traitements plus complexes que les aires primaires. Enfin, l'organisation structurale du cerveau est susceptible d'être partiellement remodelée par le phénomène de plasticité corticale, qui peut aboutir à d'importantes réorganisations du réseau neuronal (voir *« Réorganisation neurofonctionnelle après une altération sensorielle »*, ch. 8, p. 155 et *« Réorganisation neurofonctionnelle après implantation cochléaire »*, ch. 9, p. 191). Ce phénomène est étroitement associé à la notion de période critique de plasticité, qui a trait aux contraintes temporelles auxquelles est soumise la réorganisation corticale.

Chacun des hémisphères est divisé en quatre lobes anatomiquement distincts, désignés respectivement comme frontal, situé en avant, temporal, situé sur le côté, pariétal, situé au-dessus et en arrière, et occipital, situé en arrière. Les lobes frontaux sont principalement en relation avec la planification des actions et le contrôle volontaire des mouvements. Les lobes pariétaux, situés derrière les lobes frontaux, desquels ils sont séparés par un sillon profond, le sillon central, ou scissure de Rolando (*central sulcus*), sont principalement en relation avec le sens du toucher et l'image du corps. Les lobes occipitaux, situés en arrière des lobes pariétaux, sur la face arrière du cerveau, permettent le traitement des informations visuelles. Les lobes temporaux sont séparés des lobes frontaux et pariétaux par un sillon encore plus profond, le sillon latéral, ou scissure de Sylvius (*lateral sulcus*), et permettent le traitement des informations auditives mais également des informations multimodales, de la mémoire et des émotions.

A.1.2 Aires de Brodmann[2]

Les aires de Brodmann† sont les délimitations anatomiques du cerveau humain définies au début du XX^{e} siècle par Korbinian Brodmann sur la base de considérations cytoarchitectoniques. En effet l'organisation cellulaire des couches du cortex diffère selon les emplacements, ce qui permet de définir des aires partageant la même organisation cellulaire (nombre et épaisseur des couches, arborisation dendritique, etc.) Les aires de Brodmann sont très largement utilisées en neurosciences afin de permettre un repérage dans la topographie corticale.

[2] d'après Dubin (1999)

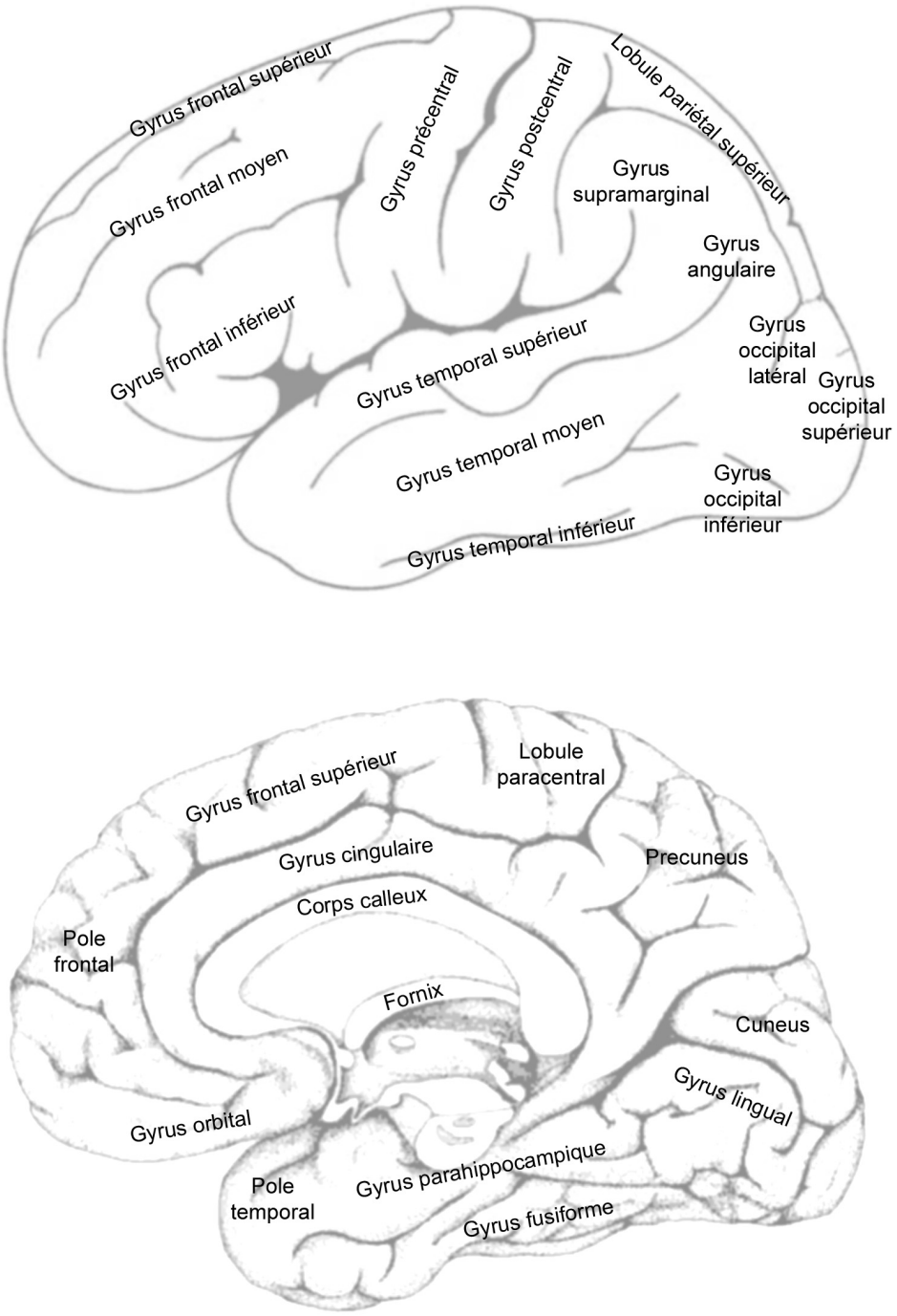


FIG. A.2 – Principaux gyri et sillons corticaux.

1 - AIRE POSTCENTRALE INTERMÉDIAIRE (*area postcentralis intermedia*). Située dans le gyrus postcentral. Délimitée cytoarchitectoniquement par les aires postcentrale rostrale 3 et postcentrale caudale 2, et à son extrémité ventrale par l'aire subcentrale 43.

2 - AIRE POSTCENTRALE CAUDALE (*area postcentralis caudalis*). Située principalement dans la partie caudale du gyrus postcentral et la lèvre rostrale du sillon postcentral avec une extension caudale le long du sillon intrapariétal. Délimitée cytoarchitectoniquement rostralement par l'aire postcentrale

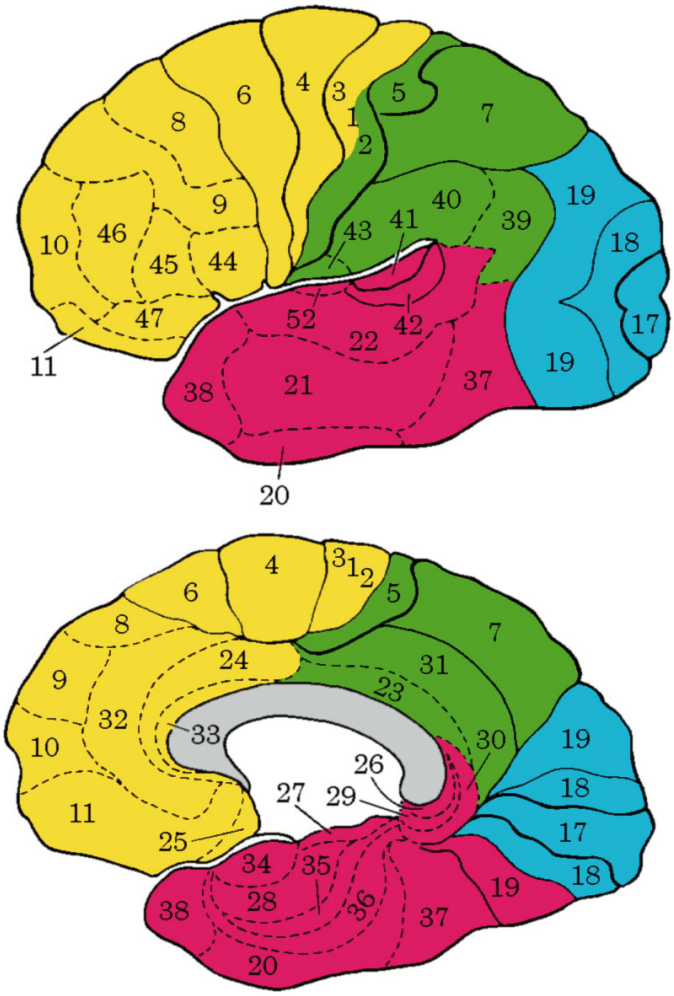


FIG. A.3 – Aires de Brodmann : vues latérale et sagittale médiale.

intermédiaire 1 et caudalement par les aires prépariétale 5, supérieure pariétale 7 et supramarginale 40.

3 - AIRE POSTCENTRALE ROSTRALE (*area postcentralis oralis*). Située principalement dans la partie rostrale du gyrus postcentral, comprenant la face caudale du sillon central. A chaque extrémité du sillon elle s'étend au-delà du creux du sillon jusqu'au gyrus précentral. Délimitée cytoarchitectoniquement rostralement par l'aire gigantopyramidale 4 et caudalement par l'aire postcentrale intermédiaire 1.

4 - AIRE GIGANTOPYRAMIDALE (*area gigantopyramidalis*). Située dans le gyrus précentral. Cytoarchitectoniquement, la frontière caudale avec l'aire précentrale rostrale 3 ne coïncide pas précisément avec le fond du sillon central mais se situe de façon variable dans les creux du gyrus postcentral et du gyrus précentral. Par ailleurs cette aire ne s'étend pas systématiquement au sillon cingulaire médialement ou au bout du sillon central ventro-latéralement. Délimitée rostralement par l'aire frontale agranulaire 6.

5 - AIRE PRÉPARIÉTALE (*area praeparietalis*). Occupe le lobule pariétal supérieur et une portion du gyrus postcentral, particulièrement sur le versant médial de l'hémisphère. Délimitée approximativement par le sillon cingulaire sur le versant médial de l'hémisphère et par le sillon postcentral supérieur sur le versant latéral. Délimitée cytoarchitectoniquement par les aires postcentrale caudale 2, supérieure pariétale 7 et sur la face médiale de l'hémisphère par les aires gigantopyramidale 4 et cingulaire postérieure dorsale 31.

6 - Aire frontale agranulaire (*area postcentralis oralis*). Située principalement dans les partie caudales du gyrus frontal supérieur et du gyrus frontal moyen et les parties rostrales du gyrus précentral inoccupées par l'aire gigantopyramidale 4. Elle s'étend du sillon cingulaire sur le versant médial de l'hémisphère au sillon latéral sur le versant latéral. Délimitée cytoarchitectoniquement rostralement par la région frontale et caudalement par l'aire gigantopyramidale 4.

7 - Aire supérieure pariétale (*area parietalis superior*). Occupe l'essentiel du lobule pariétal supérieur et une partie du precuneus. Délimitée approximativement par le sillon postcentral supérieur rostralement, le sillon intrapariétal latéralement, le sillon pariéto-occipital caudalement et, sur la face médiale de l'hémisphère, le sillon subpariétal. Délimitée cytoarchitectoniquement rostralement par les aires prépariétale 5 et postcentrale caudale 2, caudalement par l'aire préstriée 19 et médialement par l'aire cingulaire postérieure dorsale 31.

8 - Aire frontale intermédiaire (*area frontalis intermedia*). Située principalement dans le gyrus frontal supérieur, s'étendant à partir du sillon cingulaire sur la surface médiale du bord de l'hémisphère jusqu'au gyrus frontal moyen. Délimitée cytoarchitectoniquement caudalement par l'aire frontale agranulaire 6 et ventralement par l'aire frontale granulaire 9.

9 - Aire frontale granulaire (*area frontalis granulaireis*). Occupe des parties du gyrus frontal supérieur et du gyrus frontal moyen. Sa frontière approximative sur le versant médial de l'hémisphère est le sillon cingulaire et, sur le versant latéral, le sillon frontal inférieur. Délimitée cytoarchitectoniquement dorso-caudalement par l'aire frontale intermédiaire 8, caudalement par l'aire frontale agranulaire 6, et ventralement par les aires frontopolaire 10, frontale moyenne 46 et operculaire 44.

10 - Aire frontopolaire (*area frontopolaris*). Occupe l'essentiel des parties rostrales du gyrus frontal supérieur et du gyrus frontal moyen. Sur le versant médial de l'hémisphère elle est délimitée ventralement par le sillon rostral supérieur. Elle ne s'étend pas au-delà du sillon cingulaire. Délimitée cytoarchitectoniquement dorsalement par l'aire frontale granulaire 9, caudalement par l'aire frontale moyenne 46, et ventralement par les aires orbitale 47 et frontopolaire 12.

11 - Aire préfrontale (*area praefrontalis*). Constitue l'essentiel des gyri orbitaux, du gyrus droit et la partie la plus rostrale du gyrus frontal supérieur. Délimitée médialement par le sillon rostral inférieur et latéralement approximativement par le sillon frontomarginal. Délimitée cytoarchitectoniquement sur les versants rostral et latéral de l'hémisphère par les aires frontopolaire 10, orbitale 47 et triangulaire 45 ; sur la surface médiale elle est délimitée dorsalement par l'aire préfrontale 12 et caudalement par l'aire subgénuale 25.

12 - Aire préfrontale (*area praefrontalis*). Occupe l'espace entre le sillon rostral supérieur et le sillon rostral inférieur. Délimitée cytoarchitectoniquement dorsalement par l'aire frontopolaire 10 et l'aire cingulaire antérieure dorsale 32 ; caudalement, ventralement et rostralement elle est délimitée par l'aire préfrontale 11. (Originalement décrite comme une partie de l'aire préfrontale 11 mais non montrée sur la carte, par conséquent elle a été repérée comme une aire indépendante 12).

17 - Aire striée (*area striata*). Partie du lobe occipital du cortex définie sur la base de sa cyto- et de sa myeloarchitecture, principalement par les bandes de Gennari.

18 - Aire parastriée (*area parastriata*). Située dans les parties occipitales du cuneus, du gyrus lingual et du gyrus occipital latéral. Délimitée cytoarchitectoniquement sur un côté par l'aire striée 17, de laquelle elle se distingue par l'absence de bandes de Gennari, et sur l'autre par l'aire préstriée 19.

19 - Aire péristriée (*area peristriata*). Située dans les parties occipitales du gyrus lingual, du cuneus, du gyrus occipital latéral et du gyrus occipital supérieur, et délimitée approximativement par le sillon pariéto-occipital. Délimitée cytoarchitectoniquement sur un côté par l'aire parastriée 18 qu'elle entoure. Rostralement elle est délimitée par les aires angulaire 39 et occipitotemporale 37.

20 - Aire temporale inférieure (*area temporalis inferior*). Correspond approximativement au gyrus temporal inférieur. Délimitée cytoarchitectoniquement médialement par l'aire ectorhinale 36, latéralement par l'aire temporale moyenne 21, rostralement par l'aire temporopolaire 38 et caudalement par l'aire occipitotemporale 37.

21 - Aire temporale moyenne (*area temporalis media*). Correspond approximativement au gyrus temporal moyen. Délimitée rostralement par l'aire temporopolaire 38, ventralement par l'aire temporale inférieure 20, caudalement par l'aire occipitotemporale 37, et dorsalement par l'aire supérieure temporale 22.

22 - Aire temporale supérieure (*area temporalis superior*). Correspond approximativement aux deux tiers latéral et caudal du gyrus temporal supérieur. Délimitée rostralement par l'aire temporopolaire 38, médialement par l'aire temporale transverse postérieure 42, ventro-caudalement par l'aire temporale moyenne 21 et dorso-caudalement par l'aire supramarginale 39.

23 - Aire cingulaire postérieure ventrale (*area cingularis posterior ventralis*). Occupe l'essentiel du gyrus cingulaire postérieur adjacent au corps calleux. A l'extrémité caudale elle est délimitée approximativement par le sillon pariéto-occipital. Délimitée cytoarchitectoniquement dorsalement par l'aire cingulaire postérieure dorsale 31, rostralement par l'aire cingulaire antérieure ventrale 24, et ventro-rostralement dans sa moitié caudale par la région rétrospléniale.

24 - Aire cingulaire antérieure ventrale (*area cingularis anterior ventralis*). Occupe l'essentiel du gyrus cingulaire antérieur dans un arc entourant le coude du corps calleux. Sa frontière externe correspond approximativement au sillon cingulaire. Délimitée cytoarchitectoniquement intérieurement par l'aire prégénuale 33, extérieurement par l'aire cingulaire antérieure dorsale 32, et caudalement par les aires cingulaire postérieure ventrale 23 et cingulaire postérieure dorsale 31.

25 - Aire subgénuale (*area subgenualis*). Etroite bande localisée dans la partie caudale de l'aire subcalleuse adjacente au gyrus paraterminal dont elle est séparée par le sillon parolfactif postérieur. Elle est délimitée rostralement par l'aire préfrontale 11 et caudalement par le gyrus paraterminal.

26 - Aire ectosplénial (*area ectosplenialis*). Etroite bande localisée dans l'isthme du gyrus cingulaire adjacent au gyrus fasciolaire intérieurement. Elle est délimitée extérieurement par l'aire granulaire rétrolimbique 29.

28 - Aire entorhinale (*area entorhinalis*). Située dans l'aire entorhinale sur le versant médial du lobe temporal. Avec l'aire entorhinale dorsale 34, ces deux aires constituent approximativement l'aire entorhinale.

29 - Aire rétrolimbique granulaire (*area retrolimbica granulaireis*). Etroite bande située dans l'isthme du gyrus cingulaire. Délimitée cytoarchitectoniquement intérieurement par l'aire ectosplénial 26 et extérieurement par l'aire rétrolimbique agranulaire 30.

30 - Aire rétrolimbique agranulaire (*area retrolimbica agranulaireis*). Située dans l'isthme du gyrus cingulaire. Délimitée cytoarchitectoniquement intérieurement par l'aire rétrolimbique granulaire 29, dorsalement par l'aire cingulaire postérieure ventrale 23 et ventro-latéralement par l'aire ectorhinale area 36.

31 - Aire cingulaire postérieure dorsale (*area cingularis posterior dorsalis*). Occupe des parties du gyrus cingulaire postérieur et du versant médial du lobe pariétal. Ses frontières approximatives sont le sillon cingulaire dorsalement et le sillon pariéto-occipital caudalement. Elle entoure partiellement le sillon subpariétal. Délimitée cytoarchitectoniquement rostralement par l'aire cingulaire antérieure ventrale 24, ventralement par l'aire cingulaire postérieure ventrale 23, dorsalement par les aires gigantopyramidale 4 et prépariétale 5, et caudalement par l'aire supérieure pariétale 7.

32 - Aire cingulaire antérieure dorsale (*area cingularis anterior dorsalis*). Forme un arc extérieur autour du gyrus cingulaire antérieur. Le sillon cingulaire définit approximativement sa frontière intérieure et le sillon rostral supérieur sa frontière ventrale ; rostralement elle s'étend presque jusqu'à la bordure du lobe frontal. Délimitée cytoarchitectoniquement intérieurement par l'aire cingulaire antérieure ventrale 24, extérieurement par les limites médiales des aires frontale agranulaire 6, frontale intermédiaire 8, frontale granulaire 9, frontopolaire 10, et préfrontale 11.

33 - Aire prégenuale (*area praegenualis*). Etroite bande située dans le gyrus cingulaire antérieur adjacent au gyrus supracalleux dans l'intérieur du sillon calleux. Délimitée cytoarchitectoniquement par l'aire cingulaire antérieure ventrale 24 et le gyrus supracalleux.

34 - Aire dorsale entorhinale (*area entorhinalis dorsalis*). Située dans l'aire entorhinale sur le versant médial du lobe temporal. Avec l'aire entorhinale 28, ces deux aires constituent approximativement l'aire entorhinale.

35 - Aire périrhinale (*area perirhinalis*). Située le long du sillon rhinal. Cytoarchitectoniquement délimitée médialement par l'aire entorhinale 28 et latéralement par l'aire ectorhinale area 36.

36 - Aire ectorhinale (*area aire ectorhinaleis*). Située principalement dans le gyrus fusiforme, avec sa frontière médiale correspondant approximativement au sillon rhinal. Délimitée cytoarchitectoniquement latéralement et caudalement par l'aire temporale inférieure 20, médialement par l'aire périrhinale 35 et rostralement par l'aire temporopolaire 38.

37 - Aire occipitotemporale (*area occipitotemporalis*). Située principalement dans les parties caudales du gyrus fusiforme et du gyrus temporal inférieur sur les surfaces médiobasale et latérale de l'extrémité caudale du lobe temporal. Délimitée cytoarchitectoniquement caudalement par l'aire péristriée 19, rostralement par les aires temporale inférieure 20 et temporale moyenne 21, et dorsalement sur le versant latéral de l'hémisphère par l'aire angulaire 39.

38 - Aire temporopolaire (*area temporopolaris*). Située principalement dans les parties les plus rostrales du gyrus temporal supérieur et du gyrus temporal moyen. Délimitée cytoarchitectoniquement caudalement par les aires temporale inférieure 20, temporale moyenne 21, supérieure temporale 22 et ectorhinale 36.

39 - Aire angulaire (*area angularis*). Correspond à la partie du gyrus angulaire qui entoure l'extrémité caudale du sillon temporal supérieur. Dorsalement elle est délimitée approximativement par le sillon intrapariétal. Délimitée cytoarchitectoniquement rostralement par l'aire supramarginale 40, dorso-caudalement par l'aire préstriée 19, et ventralement par l'aire occipitotemporale 37.

40 - Aire supramarginale (*area supramarginalis*). Située principalement dans le gyrus supramarginal entourant la branche postérieure ascendante du sillon latéral. Délimitée approximativement par le sillon intrapariétal, le sillon postcentral inférieur, le sillon subcentral postérieur et le sillon latéral. Délimitée cytoarchitectoniquement caudalement par l'aire angulaire 39, dorso-rostralement par l'aire postcentrale caudale 2, et ventralement par les aire subcentrale 43 et supérieure temporale 22.

41 - Aire temporale transverse antérieure (*area temporalis transversa anterior*). Occupe le gyrus temporal transverse antérieur dans le creux du sillon latéral sur la surface dorsale du lobe temporal. Délimitée cytoarchitectoniquement médialement par l'aire parainsulaire 52 et latéralement par l'aire temporale transverse postérieure 42.

42 - Aire temporale transverse postérieure (*area temporalis transversa posterior*). Située dans le creux du sillon latéral sur la surface dorsale du lobe temporal. Délimitée cytoarchitectoniquement médialement par l'aire temporale transverse antérieure 41 et latéralement par l'aire supérieure temporale 22.

43 - Aire subcentrale (*area subcentralis*). Occupe le gyrus postcentral et le gyrus précentral entre l'extrémité ventro-latérale du sillon central et le fond du sillon latéral au niveau de l'insula. Ses frontières rostrale et caudale sont approximativement délimitées par, respectivement, le sillon subcentral antérieur et le sillon subcentral postérieur. Délimitée cytoarchitectoniquement rostralement par l'aire frontale agranulaire 6, et caudalement, pour l'essentiel, par les aires postcentrale caudale 2 et supramarginale 40.

44 - Aire operculaire (*area opercularis*). Correspond approximativement à la partie operculaire du gyrus frontal inférieur. Délimitée caudalement par le sillon précentral inférieur et rostralement par la branche ascendante antérieure du sillon latéral. Elle entoure le sillon diagonal. Dans le fond du sillon latéral elle est délimitée par l'insula. Cytoarchitectoniquement délimitée dorso-caudalement par l'aire frontale agranulaire 6, dorsalement par l'aire frontale granulaire 9 et rostralement par l'aire triangulaire 45.

45 - Aire triangulaire (*area triangularis*). Occupe la partie triangulaire du gyrus frontal inférieur et, entourant la branche horizontale antérieure du sillon latéral, une portion de la partie orbitale du gyrus frontal inférieur. Délimitée caudalement par la branche ascendante antérieure du sillon latéral, elle s'arrête sur l'insula au fond du sillon latéral. Cytoarchitectoniquement délimitée caudalement par l'aire operculaire 44, dorso-rostralement par l'aire frontale moyenne 46 et ventralement par l'aire orbitale 47.

46 - Aire frontale moyenne (*area frontalis media*). Occupe approximativement le tiers moyen du gyrus frontal moyen et la partie la plus rostrale du gyrus frontal inférieur. Délimitée cytoarchitectoniquement dorsalement par l'aire frontale granulaire 9, ventro-rostralement par l'aire frontopolaire 10 et caudalement par l'aire triangulaire 45.

47 - Aire orbitale (*area orbitalis*). Entoure la partie caudale du sillon orbital à partir de laquelle elle s'étend latéralement dans la partie orbitale du gyrus frontal inférieur. Cytoarchitectoniquement délimitée caudalement par l'aire triangulaire 45, médialement par l'aire préfrontale 11, et rostralement par l'aire frontopolaire 10.

48 - Aire retrosubiculaire (*area retrosubicularis*). Située sur la surface médiale du lobe temporal. Cytoarchitectoniquement délimitée rostralement par l'aire périrhinale 35 et médialement par le présubiculum. (Décrite par Brodmann mais non inclue dans ses cartes du cortex humain).

52 - Aire parainsulaire (*area parainsularis*). Située dans le creux du sillon latéral sur la surface dorsale du lobe temporal. Sa frontière médiale correspond approximativement à la jonction entre le lobe temporal et l'insula. Délimitée cytoarchitectoniquement latéralement par l'aire temporale transverse antérieure 41.

A.2 Normalisation de l'anatomie corticale

A.2.1 Variabilité anatomique

Les sulci et les gyri permettent une subdivision topographique naturelle de l'anatomie corticale. Par ailleurs, les zones où se rejoignent les régions définies par leur structure architectonique ou leur rôle fonctionnel sont fréquemment situées le long du lit des sulci majeurs ou mineurs du cortex (Rademacher *et al.*, 1993; Roland et Zilles, 1994). Une considérable variabilité de la taille, de la forme et de la configuration des gyri et sulci corticaux a été démontrée, cette variabilité devenant particulièrement évidente lors de l'analyse de la surface du cortex (Evans *et al.*, 1992; Mazziotta *et al.*, 1995; Zilles *et al.*, 1997; Mazziotta *et al.*, 2001). On peut observer ces différences anatomiques non seulement selon le sexe et l'âge, mais également en fonction du côté observé (Steinmetz *et al.*, 1991; Witelson et Kigar, 1992; Kulynych *et al.*, 1994; Sowell *et al.*, 1999; Kansaku *et al.*, 2000; Sowell *et al.*, 2002).

A.2.2 Normalisation spatiale

La normalisation spatiale est une méthode très utilisée pour comparer entre elles les images cérébrales. Des méthodes puissantes d'ajustement non-linéaire ont été développées dans ce but. La plupart se basent sur la maximisation d'une mesure de similarité entre l'image à traiter et une image de référence. Ainsi la normalisation spatiale est rendue possible par l'utilisation de procédures comme celle implémentée dans SPM (statistical parametric mapping)† (Friston *et al.*, 1995; Ashburner et Friston, 1999; Functional Imaging Laboratory, 2007).

Bien que ces méthodes soient très efficaces pour corriger des variations globales dans la forme du cortex, leur comportement local au niveau des plis corticaux est très variable. Par ailleurs, l'absence d'homologies clairement définies entre les sillons corticaux individuels empêche l'établissement de normes standard pour l'ajustement de la forme corticale (Mangin *et al.*, 2004). Les recherches les plus récentes en neuroimagerie fonctionnelle ont permis de développer des outils d'ajustement locaux au niveau du pli cortical (Mangin *et al.*, 2004). L'utilisation progressive à grande échelle de ces outils est d'une importance primordiale pour la précision et l'adéquation des résultats issus de l'imagerie neurofonctionnelle.

A.2.3 Atlas de référence

L'atlas cérébral de référence utilisé par les procédures de normalisation anatomique dans SPM est un cerveau moyenné à partir de 152 cerveaux individuels acquis par IRM, créé par l'Institut neurologique de Montréal MNI (Montreal Neurological Institute)†. Cette image de référence est connue sous l'appellation *MNI average* (Collins *et al.*, 1994). Toutefois, en raison de l'absence de caractéristiques anatomiques détaillées sur l'image moyennée, les résultats d'imagerie fonctionnelle sont la plupart du temps surimposés à une image IRM haute résolution pondérée en T1, désignée par *MNI single-subject*. Comme cette image n'est pas étiquetée anatomiquement, la plupart des chercheurs préfèrent recourir à l'atlas de Talairach et Tournoux† (Talairach et Tournoux, 1988) afin de rapporter la localisation des activations issues des études en neuroimagerie fonctionnelle, car cet atlas fournit une description anatomique détaillée dans l'espace stéréotaxique, incluant la subdivision en aires de Brodmann. Cet atlas n'est cependant pas exempt de défauts, le cerveau utilisé étant légèrement atrophié.

Bibliographie

Altenmuller, E. (2003). Focal dystonia : advances in brain imaging and understanding of fine motor control in musicians. *Hand. Clin.*, 19(3):523–538, xi.

Ashburner, J. et Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Hum. Brain Mapp.*, 7:254–266.

Collins, D. L., Neelin, P., Peters, T. M. et Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.*, 18:192–205.

Dubin, M. (1999). Locational descriptions of human brodmann areas. Site internet. http://spot.colorado.edu/~dubin/talks/brodmann/neuronames.html.

Evans, A. C., Marrett, S., Neelin, P., Collins, L., Worsley, K., Dai, W., Milot, S., Meyer, E. et Bub, D. (1992). Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage*, 1(1):43–53.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D. et Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging : A general linear approach. *Hum. Brain Mapp.*, 2:189–210.

Functional Imaging Laboratory (2007). Statistical parametric mapping. Site internet. http://www.fil.ion.ucl.ac.uk/spm/.

Kansaku, K., Yamaura, A. et Kitazawa, S. (2000). Sex differences in lateralization revealed in the posterior language areas. *Cereb. Cortex*, 10(9):866–872.

Kulynych, J. J., Vladar, K., Jones, D. W. et Weinberger, D. R. (1994). Gender differences in the normal lateralization of the supratemporal cortex : MRI surface-rendering morphometry of Heschl's gyrus and the planum temporale. *Cereb. Cortex*, 4(2):107–118.

Mangin, J. F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F. et Regis, J. (2004). A framework to study the cortical folding patterns. *Neuroimage*, 23:1.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R. et Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain : International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 356(1412):1293–1322.

Mazziotta, J. C., Toga, A. W., Evans, A., Fox, P. et Lancaster, J. (1995). A probabilistic atlas of the human brain : theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). *Neuroimage*, 2(2):89–101.

Rademacher, J., Caviness, Jr, V. S., Steinmetz, H. et Galaburda, A. M. (1993). Topographical variation of the human primary cortices : implications for neuroimaging, brain mapping, and neurobiology. *Cereb. Cortex*, 3(4):313–329.

Roland, P. E. et Zilles, K. (1994). Brain atlases–a new research tool. *Trends Neurosci.*, 17(11):458–467.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Batth, R., Jernigan, T. L. et Toga, A. W. (1999). Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *Neuroimage*, 9(6 Pt 1):587–597.

Sowell, E. R., Thompson, P. M., Rex, D., Kornsand, D., Tessner, K. D., Jernigan, T. L. et Toga, A. W. (2002). Mapping sulcal pattern asymmetry and local cortical surface gray matter distribution in vivo : maturation in perisylvian cortices. *Cereb. Cortex*, 12(1):17–26.

Steinmetz, H., Volkmann, J., Jancke, L. et Freund, H. J. (1991). Anatomical left-right asymmetry of language-related temporal cortex is different in left- and right-handers. *Ann. Neurol.*, 29(3):315–319.

Talairach, J. et Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain 3-Dimensional Proportional System : An Approach to Cerebral Imaging*. Stuttgart, New York.

Witelson, S. F. et Kigar, D. L. (1992). Sylvian fissure morphology and asymmetry in men and women : bilateral differences in relation to handedness in men. *J. Comp. Neurol.*, 323(3):326–340.

Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., Schormann, T., Mohlberg, H., Bürgel, U., Steinmetz, H., Schlaug, G. et Roland, P. E. (1997). Quantitative analysis of sulci in human cerebral cortex : development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. *Hum. Brain Mapp.*, 5:218–221.

Annexe B

Publications

Sommaire

Rouger et al. (2007). Proc. Natl. Acad. Sci. U.S.A. 104, 7295–7300. . . . 353
Rouger et al. (2008). Brain Res. 1188, 87–99. 359
Strelnikov et al. (2009). Neuropsychologia 47, 972–979. 373

SEE COMMENTARY

Evidence that cochlear-implanted deaf patients are better multisensory integrators

J. Rouger*, S. Lagleyre†, B. Fraysse†, S. Deneve‡, O. Deguine*†, and P. Barone*§

*Centre de Recherche Cerveau et Cognition, Université Toulouse 3, Centre National de la Recherche Scientifique, Faculté de Médecine de Rangueil, 31062 Toulouse Cedex 9, France; †Service d'Oto–Rhino–Laryngologie, Hopital Purpan, FR-31059 Toulouse, France; and ‡Groupe de Neurosciences Théoriques, École Normale Supérieure, 3, Rue d'Ulm, 75005 Paris, France

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved February 26, 2007 (received for review October 24, 2006)

The cochlear implant (CI) is a neuroprosthesis that allows profoundly deaf patients to recover speech intelligibility. This recovery goes through long-term adaptative processes to build coherent percepts from the coarse information delivered by the implant. Here we analyzed the longitudinal postimplantation evolution of word recognition in a large sample of CI users in unisensory (visual or auditory) and bisensory (visuoauditory) conditions. We found that, despite considerable recovery of auditory performance during the first year postimplantation, CI patients maintain a much higher level of word recognition in speechreading conditions compared with normally hearing subjects, even several years after implantation. Consequently, we show that CI users present higher visuoauditory performance when compared with normally hearing subjects with similar auditory stimuli. This better performance is not only due to greater speechreading performance, but, most importantly, also due to a greater capacity to integrate visual input with the distorted speech signal. Our results suggest that these behavioral changes in CI users might be mediated by a reorganization of the cortical network involved in speech recognition that favors a more specific involvement of visual areas. Furthermore, they provide crucial indications to guide the rehabilitation of CI patients by using visually oriented therapeutic strategies.

cochlear implant | deafness | multisensory integration | speech comprehension

Despite the apparent division between sensory modalities from the receptors to high cortical levels, we can simultaneously integrate visual and auditory signals resulting in qualitative percepts distinct from those derived from a single unisensory stimulus (1, 2). Furthermore, in cases of precise temporal or spatial congruency between the bisensory stimuli, multisensory integration is expressed at the behavioral level by perceptual improvements by reducing ambiguity (3, 4) and at the neuronal level by enhancing neuronal activity (5). Multisensory integration is also essential for speech recognition, which is based on the simultaneous integration of visual information derived from lip movements and auditory cues produced by the talker (6). The McGurk effect, in which a mismatch between the visual and auditory speech signals is artificially introduced, reveals that the visual information derived from lip movements can strongly influence our auditory perception (7). Although we might not be aware of the relevance of the visual cues for normal speech recognition, the influence of vision becomes convincingly apparent when the auditory information is embedded in noise. In degraded auditory conditions, the visuoauditory presentation leads to higher performance of recognition, when compared with the auditory alone stimulation (8, 9), in a mechanism that mimics an improvement in the acoustic signal-to-noise ratio (SNR) (10).

In normally hearing (NH) subjects, although speechreading performance is very low, the association during development between the auditory and visual speech information is critical for a normal acquisition of multisensory speech perception (11). Speechreading can become extremely crucial in the case of profound deafness because the acquisition of strong skills in speechreading is one of the sensory substitution strategies developed by deaf patients to access speech recognition. During recent decades, remarkable technical progress has been made in the efficiency of cochlear implants (CIs), allowing patients afflicted with sensorineural hearing loss to recover a large range of auditory functions (12–14). Although in postlinguistically deafened adults the implantation of a cochlear neuroprosthesis allows a significant recovery of auditory speech intelligibility (13, 15), CI patients remain highly sensitive to noisy environments and have impaired speech-recognition performance in the presence of masking sounds (16). This suggests that CI users might develop specific visual and visuoauditory integration skills to overcome their difficulties in speech recognition in everyday life. Although it is known that deaf people develop specific visual abilities (17, 18), including speechreading (19), no data are available on the postimplantation evolution of visual performance in CI users and how these abilities are related to the recovery of the auditory functions. To answer these questions, we have analyzed the longitudinal postimplantation performance of CI patients in word recognition presented in a uni- or bisensory modality. For further analysis, we compared CI users' performance to that of NH subjects who were presented with identical stimuli, with an additional sound computation designed to simulate the processing performed by a CI system. These behavioral data, combined with a multisensory integration model, reveal that, as a consequence of their auditory deprivation, CI deaf patients have developed a much higher proficiency to fuse visual speech information to auditory cues as compared with NH subjects.

Results

Pre- and Postimplantation Performance of Speech Recognition in CI Patients. We have analyzed the performance of 97 CI users in disyllabic word recognition by using three modalities (auditory, visual, and visuoauditory) during a longitudinal study that extended over 8 years after implantation (Fig. 1).

Auditory Speech. First, at the time the implant is switched on (T0), CI users obtain a significant recovery of word recognition in auditory modality, with a performance level of 47.1 ± 27.3% SD in quiet conditions. This performance level is much higher than that obtained before implantation by using an external hearing aid (mean 10.4 ± 14.2% correct, $P < 0.05$). Auditory performance increases significantly during the subsequent months ($P < 0.05$), before reaching a plateau from about the seventh month on and then showing no significant improvement in the following years (mean 81% over the first year).

Author contributions: J.R. and S.L. contributed equally to this work; J.R., S.L., B.F., S.D., O.D., and P.B. designed research; J.R., S.L., B.F., O.D., and P.B. performed research; J.R., S.L., S.D., and P.B. analyzed data; and J.R., S.L., S.D., O.D., and P.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: CI, cochlear implant; NH, normally hearing; SNR, signal-to-noise ratio.

See Commentary on page 6883.

§To whom correspondence should be addressed. E-mail: pascal.barone@cerco.ups-tlse.fr.

© 2007 by The National Academy of Sciences of the USA

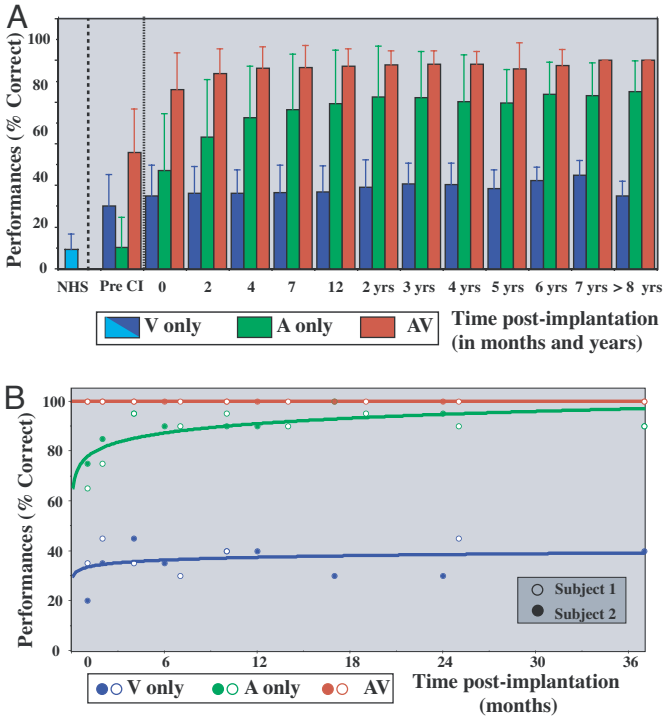


Fig. 1. Word-recognition scores for CI users in the three sensory modalities: auditory only (A only, green), visual only (V only, blue), and bisensory visuoauditory (VA, red). (*A*) Longitudinal performance (mean percentage correct ± SD) of the entire cohort of CI users ($n = 97$) at different times before (preoperatively) and after the cochlear implantation. In the left part of the graph, we have reported the speechreading performance (V only) of NH subjects ($n = 42$). (*B*) Individual performance levels for two CI users who have been regularly evaluated during a 3-year period after implantation. Both graphs show the significant recovery of auditory speech recognition during the first year compared with the weak performance before implantation and the stable high values of speechreading at all periods tested. In bisensory conditions, near-maximum scores are achieved.

Speechreading. At T0, speechreading performance in CI users is elevated and impressively higher than that observed for NH subjects tested with the same talker (35.1 ± 14.7% vs. 9.4 ± 7.1%, respectively; $P < 0.05$). This speechreading ability in CI users is similar to that obtained a few months before implantation (mean 30.1 ± 15.1%; paired test, $P = 0.62$) and remains unchanged across all postimplantation periods tested (>35%; $P > 0.05$), although CI users have reached their maximal auditory performance. Furthermore, visual performance at T0 is not correlated with auditory proficiency ($r^2 = 0.001, P = 0.76$). At the time of implantation, duration of deafness is not correlated to visual performance ($r^2 = 0.001$, $P = 0.77$). This latter result should be tempered because most of the patients were suffering from a progressive hearing impairment, such that deafness duration could hardly be reliably defined. However, in three CI users who were affected by sudden deafness (such as meningitis) and implanted only 1 year later, speechreading performance levels were similar to that of the CI population (20%, 30%, and 45%, respectively). To strengthen this observation, we have included in our analysis data obtained from five supplementary CI users (not included in the longitudinal retrospective study) suffering from sudden deafness occurring within <1 year of implantation. In this enlarged sample ($n = 8$), we observed that performance in visual-only conditions was much higher than that observed in NH subjects (mean 27.5 ± 10.7% vs. 9.4%, respectively; $P < 0.05$). In these CI users, several months or years of auditory recovery postimplantation (auditory-only performance >90% correct) did not affect their speechreading performance. Despite the limited number of observations, this suggests that a high level of speechreading ability can be acquired rapidly during a period of auditory deprivation and then remain at stable values.

Audiovisual Speech. As expected from the classically perceptual benefit derived from multisensory integration (5), prior to implantation, CI users present higher performance in visuoauditory conditions compared with the auditory-alone conditions (55.8 ± 21.0% vs. 10.4%, respectively; $P < 0.05$). A similar effect is observed in CI users postimplantation; when compared with unisensory conditions, visuoauditory integration results in an improvement in word recognition in CI patients tested at T0 (86 ± 17.4% correct; $P < 0.05$ for both comparisons). From the time of implantation, audiovisual recognition improved slightly ($P < 0.05$) with practice, allowing CI users to reach near-perfect performance levels (94 ± 12.0%) as early as the second month postimplantation.

We believe that the difference in bisensory performance of CI users when comparing pre- and postimplantation periods (55.8% vs. 86% at T0; $P < 0.05$) is derived from higher auditory performance provided by the neuroprosthesis. In agreement with this, we saw that, in a limited number of CI users ($n = 14$) who did not show any improvement in auditory word recognition, the visuoauditory gain remains unchanged when comparing the two testing periods (mean visuoauditory benefit preimplantation 0.54 vs. 0.62 at T0; paired test, $P = 0.23$).

Comparison of Performance of NH and CI Subjects. Our results show that, during the period of deafness, CI patients have developed a specific ability in speechreading that distinguishes them from the poor speechreading skills of NH listeners. We hypothesized that this high visual aptitude might induce in CI users an improvement of the mechanisms of multisensory integration, leading to greater visuoauditory benefits than those observed in NH subjects. In this scheme, we compared the visuoauditory gain in CI users at T0 (i.e., without training) to the one obtained in naïve NH subjects exposed to a degraded auditory signal. This auditory degradation allows us to make direct comparisons of visuoauditory performance from both groups at equivalent ranges of nonoptimal auditory performance. To degrade the auditory performance of NH subjects, we first used a masking paradigm with white noise at different SNRs. In these protocols, we observed a higher recognition rate in visuoauditory versus auditory-only conditions (Fig. 2*A*), especially at intermediate SNRs at ≈15 dB (20). Second, we used a noise-band vocoder paradigm with different frequency bands that simulates the processing strategy of CIs (21). In this simulation, the global temporal and spectral information of the signal are preserved, whereas the fine temporal cues within each spectral component are removed. In this case, performance (auditory and visuoauditory) decreases rapidly as the number of bands decreases, leading to near-zero values in the two-electrode simulation (1.5% mean recognition in auditory-only presentation). However, whereas bisensory presentation improved NH subjects' performance (Fig. 2*A*), the visuoauditory gain was much lower than the one obtained in the masking protocol at equivalent auditory performance levels, suggesting that visuoauditory integration mechanisms of speech perception strongly depend on the integration of fine spectrotemporal auditory information. This hypothesis was confirmed by our model (see *Are CI Patients Better Multisensory Integrators*). When visuoauditory performance of CI users is compared with that of the NHs exposed to degraded auditory stimuli, we show that the visuoauditory gain in CI patients is higher than that observed in NH subjects in the simulation or noise-masked conditions (both comparisons, $P <$ 0.001; Fig. 2*B*). The differences in favor of CI users are especially

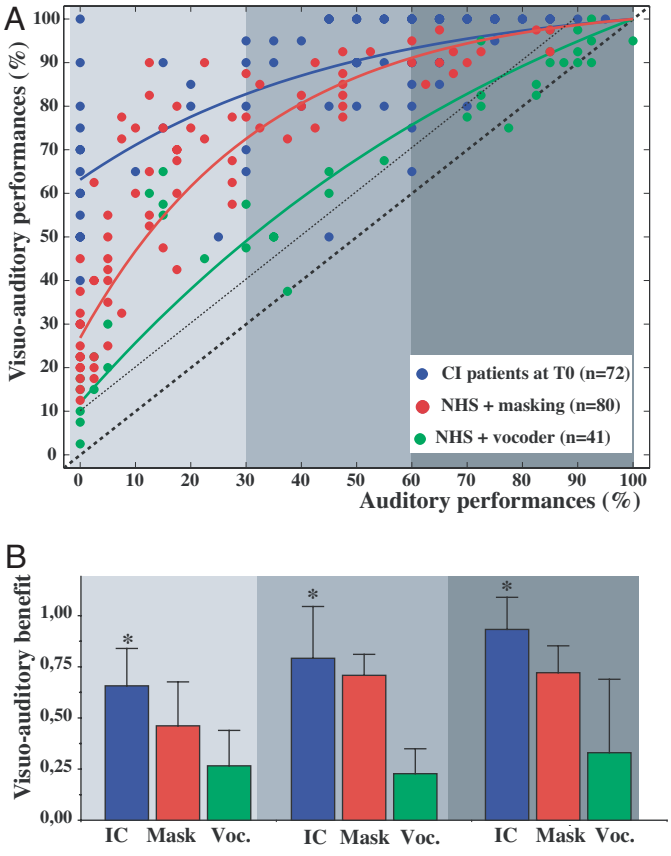


Fig. 2. Relationships between auditory and visuoauditory performance and bisensory gain. (*A*) For each group we have plotted the performance of individual subjects in auditory-only conditions with respect to the performance in visuoauditory conditions. Each point corresponds to a single subject tested in a single condition. This graph shows that the three experimental groups (CI users at T0, NH subjects with masking noise, and NH subjects with vocoder simulations) are clearly segregated in the correlation graph, indicating that the three populations present different gains after a bisensory presentation. (*B*) The visuoauditory gain (expressed as [(VA-A)/(100-A)]) is presented separately for three ranges of auditory performance (<30% correct, between 30% and 60% correct, and >60% correct) from each group. This subdivision shows that, in all cases, CI patients obtained a significantly higher visuoauditory gain than that observed in NH subjects tested with degraded auditory stimuli. The visuoauditory benefit for CI users is considerably greater than that computed from the NH subject vocoder group. The gains from these latter groups are directly comparable because of a similar processing of the auditory stimulus. Asterisks indicate statistically significant differences in visuoauditory benefits between CI users and both NH subject groups.

large in conditions of low auditory performance, where the range of correct recognition falls to <30% (CIPs vs. NHs, $P < 0.01$ for both comparisons). For example, a subset of CI users ($n = 13$) unable to perform auditory identification at all (0% correct) showed a high level of performance in the visuoauditory condition (mean 63% correct). In contrast and compared with this subset of CI users, NH subjects showing a similar level of auditory word recognition (0% correct, $n = 19$) due to highly degraded auditory conditions never reached the visuoauditory performance levels (mean 25.4% and 12.5% in masking or vocoder simulation protocols, respectively).

Although in CI patients the high efficiency of bisensory word recognition was not correlated to the level of speechreading ($r^2 = 0.068$, $P < 0.05$; see ref. 22), we further tested the hypothesis that the difference between CI and NH bisensory integration could be due to differences in absolute levels of visual performance. Consequently, we selected a subgroup of CI users showing low visual performance (lower than 20%; $n = 15$). We found within this group that the visuoauditory gain was still higher than that of NHs engaged in the simulating protocol (0.52 vs. 0.26, $P < 0.001$). In our opinion, this reinforces our conclusion that CI users have acquired a higher bisensory proficiency per se compared with NH subjects.

Are CI Patients Better Multisensory Integrators? As mentioned previously, the better performance levels of CI users compared with NHs with simulated implants could be due either to their stronger visual performance or to a better capacity for integrating visuoauditory inputs. Furthermore, electrophysiological studies have challenged whether the rules governing neuronal computing during multisensory interactions are superadditive, additive, or subadditive (23). Does it apply to the performance of speech recognition in bisensory conditions? To evaluate these hypotheses and quantify the multisensory performance, we designed two simple models of word recognition. The first model is a minimal-integration model, in the sense that the integration of auditory and visual cues occurs within the lowest possible level of interaction between both inputs (i.e., probabilistic combination). The second model is an optimal-integration model in which individual spatio- and spectrotemporal audiovisual cues are combined across modalities to minimize the amount of information required for correct word recognition. We fitted a model of optimal multisensory integration to the performance of NHs with masked auditory input (Fig. 3*D*). We then compared the performance of the model with all subjects' performance in two other conditions (CI users at T0 and NH subjects with vocoder; Fig. 3 *A* and *C*). We found that the model fitted very well the performance of CI patients, indicating that at T0 they integrate visuoauditory inputs as efficiently as NHs when their auditory input is degraded by white noise. However, the bisensory performance of NHs with simulated implants was far below the model performance levels (Fig. 3*C*). Thus, in contrast with CI users, NH subjects did not integrate their visuoauditory input optimally when this auditory input is lacking fine spectrotemporal structure. Furthermore, CI users tested 1 year postimplantation showed a significant improvement of both auditory and visuoauditory performance while keeping a constant high speechreading recognition level. When applying the model to the unisensory performance of CI users at 1 year (Fig. 3*B*), we found that the evolution of multisensory performance with practice could be entirely explained by their increased auditory performance. This finding suggests that, whereas visual and auditory inputs are integrated optimally from the start, a reorganization of auditory cortices, supporting a better capacity for dealing with distorted auditory inputs, is the main cause for the quasiperfect multisensory performance reached by CI users after 1 year.

Discussion

This study provides a long-term evaluation that shows the impressive benefits of cochlear implantation regarding the recovery of speech recognition because profoundly deaf patients can reach high rates of performance for hearing speech during the first 6 months postimplantation. The present data confirm that a profound hearing loss induces the acquisition of strong speechreading abilities (6, 19, 24, 25), but they represent the first evidence that this skill remains unaffected by the recovery of the auditory functions provided by the neuroprosthesis. CI patients preserve a striking speechreading ability acquired during the period of deafness while they have reached optimal auditory recognition. We interpret this apparently paradoxical strategy developed by CI users as a strategy to maintain through the mechanisms of bisensory integration a high level of speech recognition in a disturbed noisy auditory environment. Previous studies have reported that the performance of CI patients is highly susceptible to noise (16, 26), which is probably due to the

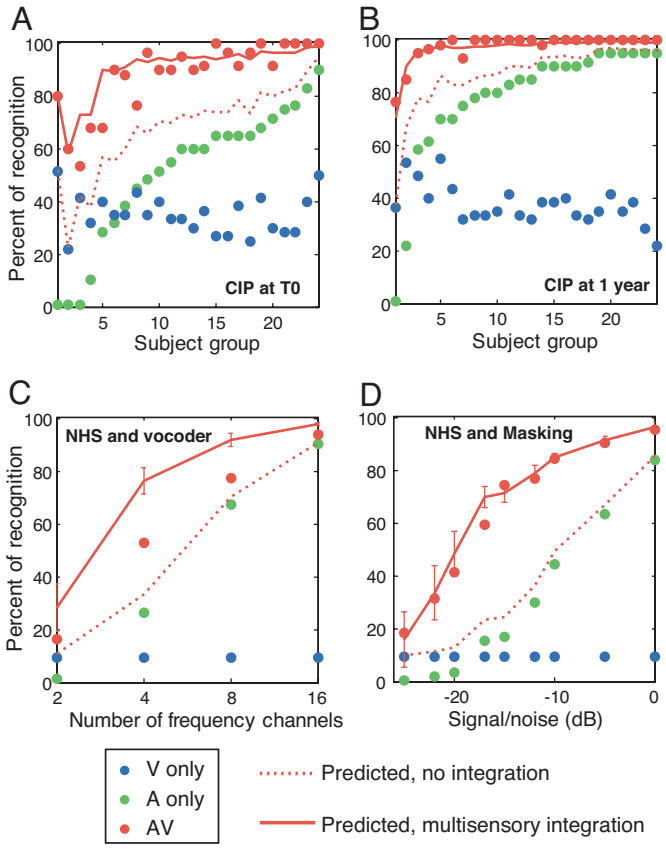


Fig. 3. Fitting a model of multisensory integration to the data. Filled circles represent the data (green, auditory; blue, visual; red, bisensory), dotted lines represent the predicted bisensory performance in the absence of multisensory integration (''model 1'' in *Materials and Methods*), and plain lines represent the prediction of an optimal multisensory integration model with a threshold $T = 6$ (''model 2'' in *Materials and Methods*). (*A* and *B*) CI users performance at T0 (*A*) and 1 year after implantation (*B*). Each group represents the averaged performance of three CI users ranked by their auditory performance. Thus, group 1 represents the (true and predicted) performance of the three subjects with the worst auditory performance, and group 26 represent the performance of the three subjects with the best auditory performance. (*C*) NH subjects' word recognition scores as a function of the amplitude of auditory white noise. (*D*) Word recognition scores of NH subjects as a function of the number of auditory frequency channels (simulated implants). This analysis shows that visuoauditory performance of CI users is perfectly in agreement with the optimal multisensory model,whereas NH subjects listening to CI simulations obtain bisensory performance far below that predicted by the integrating model.

lack of fine spectrotemporal information provided by the CI (16). In addition, as a consequence of multisensory integration (27), we show that speech intelligibility of CI users is greatly improved in visuoauditory conditions, especially during the first months postimplantation when auditory performance has not yet reached a quasioptimal level. The analysis of the performance of NH subjects submitted to degraded auditory stimuli and CI users confirmed that bisensory integration of visual and auditory speech information leads to improvement in speech intelligibility (28). However, the perceptual bisensory gain in NH subjects does not reach the same level as for CI users when compared at equivalent auditory performance. Although CI users are able to integrate their visuoauditory signal efficiently and compensate for the loss of spectral information, none of the naïve NH subjects listening to CI stimulations reach the same level of VA supraadditive integration. Altogether, we suggest that CI users have developed specific visuoauditory skills that lead to a powerful utilization of the visual spatiotemporal cues (29) provided by the lip and face movements (10), allowing these patients to reach near-perfect performance in visuoauditory situations. Using our computational model that allows us to avoid ceiling effects in subjects' performance, we confirmed that the performance of CI patients derived not only from higher efficiency in speechreading, but also from the acquisition of a higher skill level in multisensory integration when visual speech information is matched to an impoverished auditory signal.

Our results provide crucial information on the temporal window during which plastic changes can occur in the cortical network of CI patients during adaptation to the neuroprosthesis. There is now a growing body of evidence showing that sensory deprivation from early developmental stages has an important effect on the remaining sensory modalities (30, 31) through active cross-modal neuroplastic mechanisms (32, 33). In general, sensory deprivation leads to a compensatory increase in specific skills of the spared modalities that can be observed at both behavioral and neural levels in animal and human subjects (34, 35). However, cochlear implantation constitutes a unique approach to understand the cortical mechanisms that underlie the functional recuperation of the lost sensory modality. First, it has been shown that because CIs provide only a degraded signal that requires specific compensatory strategies, CI users present different levels of activation in auditory areas involved in semantic and/or phonological speech processing (36). Our longitudinal study on a large sample of patients suggests that such changes probably occur during the first 6 months depending on subjects' performance in speech recognition (37–39) and might remain different from normally hearing listeners (40) even after several years of auditory function recovery. Second, our results highlight that CI users develop a strong visuoauditory perceptive strategy for speech intelligibility while experiencing the reduced spectrotemporal information provided by the implant. This adaptation extends over the first 3 months postimplantation before being stabilized, suggesting that the pattern of brain activity during bisensory speech processing in CI users may vary during the corresponding period. Brain-imaging studies in CI deaf subjects have revealed a particular involvement of the low-level visual areas when listening to words (37). This finding corroborates our results of a strong synergy between visual and auditory processing for speech recognition following cochlear implantation. These results, in agreement with our ongoing functional imaging study (41), suggest that the visual activity derived from speechreading could actively influence the activity of the cortical network involved in hearing speech recognition and could participate in the improvement of performance in bisensory conditions. The existence of heteromodal connections that link directly unisensory areas in adult primates (42, 43) provides a possible anatomical framework for such direct visuoauditory interactions at low levels of sensory processing (44).

First, at a theoretical level, it has been shown that the fine spectrotemporal auditory information provides important cues for auditory speech recognition (45, 46). Our results broaden the role of the temporal fine structure because it optimizes the audiovisual speech integration leading to a higher multisensory perceptual benefit, in agreement with the actual technological challenge aiming at improving the spectral resolution of CIs. Second, from the clinical point of view, this work provides important cues to adapt the rehabilitation strategy as a function of implant experience. The supranormal skills in multisensory integration observed in CI deaf patients should be used to improve recovery of other auditory functions that are still deficient in CI users. Because visuoauditory training facilitates perceptual learning in a single modality (47, 48), we believe that a strong visually and audiovisually based rehabilitation during the first months postimplantation would significantly improve and fasten the functional recovery of speech intelligibility or

sound localization, which is largely deficient in unilateral CI patients (49, 50).

Materials and Methods

Participants. Our study was based on a retrospective analysis of speech recognition in 97 postlinguistically deafened subjects (mean age 56 years, range 19–82) that received a CI after profound deafness (defined as a hearing loss of ≥90 dB) of diverse etiologies (meningitis, chronic otitis, otosclerosis, neurinoma) and durations (mean age 22 years, range 1–57). The clinical implantation criteria included word and open-set sentence auditory-recognition scores <30% under best-aided conditions (i.e., with conventional acoustic hearing aids). All CI patients were recipients of a Nucleus (Cochlear) implant (CI-22 or CI-24) and used a range of different sound-coding strategies. Performance was collected during regular visits to the ear-nose-throat department following a standard rehabilitation program. We restricted our analysis to evaluations performed by the same speech therapist during the 10 years of follow-up and by using exactly the same procedures. First, we collected the performance of CI users tested before the cochlear implantation and by using an external hearing aid. On average, these tests were performed ≈6 months before the implantation (mean 5.8 months), but on 36 CI users of 97, word-recognition performance was obtained during the last 3 months preimplantation. Then from the day the CI was switched on (T0, usually 1 month postsurgery), CI users were tested at regular intervals during the first year and >8 years postimplantation. Data have been pooled into 12 groups (Fig. 1*A*) corresponding to the testing period from T0 ($n = 91$), 3 months ($n = 91$), 5 months ($n = 82$), 7 months ($n = 77$), 1 year ($n = 78$), 2 years ($n = 69$), 3 years ($n = 41$), 4 years ($n = 26$) 5 years ($n = 17$), 6 years ($n = 11$), 7 years ($n = 5$), and >8 years ($n = 4$) postimplantation. On average, CI users were tested during a period of 33 months postimplantation (± 25), with on average eight sessions per subject (±3). In this postlinguistic deaf adult population, we did not find a relationship between the age of implantation and the performance collected in the uni- or bisensory conditions (all cases $P > 0.05$). In addition, speech recognition for different sensory modalities was tested in a sample of 163 NH subjects. These control subjects were all native French speakers with self-reported normal or corrected-to-normal vision and without any previously known language or cognitive disorders.

Procedures and Stimuli. All subjects were tested on open-set recognition for French disyllabic words obtained from the classically used French speech therapist list developed by Fournier and presented in visual-only (speechreading, V), auditory-only (A), and visuoauditory (VA) conditions. Only words correctly repeated verbally by subjects were treated as correct responses (% correct score). We calculated the visual contribution to speech recognition by using the method of Sumby and Polack (8) (VA benefit = [(VA−A)/(100−A)]) to normalize for the performance observed in the A condition and thus to be able to compare directly the visuoauditory gain across groups (19). CI users were tested in silence on 20 words in each condition. In NH subjects, we developed three paradigms during which the auditory stimuli (the words pronounced by the speech therapist and recorded onto a PC computer) were differently degraded or presented without alteration. In a masking protocol, in A and VA conditions we additively combined each sound to a masking sound, with the words' acoustic level shifted to obtain the required SNR (nine SNR conditions: 0, −5, −10, −12, −15, −17, −20, −22, and −25 dB). The mask was a white noise delivered by a pseudorandom number generator and temporally modulated by monoperiodic sinusoidal lobes (period = 20 msec), with a mean rate of 300 modulations per sec temporally randomly distributed. Gain was 1 at the edge of the lobes and 0.4 at the center. This white noise modulation was carried out to ensure high random temporal fluctuations. Subjects ($n = 80$) were tested with four lists of 20 words in A or VA conditions, with masking noise at a single SNR condition and the orders for each A or VA sequence being randomized across subjects. At each masking conditions in the range from −5 and −22 SNR, data were obtained from a sample of 10 subjects (but only 3 subjects at SNR 0 dB and 7 at SNR −25 dB). In a "simulating" protocol used with a second group of NH subjects ($n = 41$), for A and VA conditions, we developed noise-band vocoder methods that simulate the signal processing computed in a CI (21). The sound was analyzed through 2, 4, 8, or 16 frequency bands by using sixth-order IIR elliptical analysis filters. The cutoff frequencies of these bands were calculated to ensure equidistance of the corresponding basilar membrane locations of the cochlea according to the human cochlear tonotopic map (51). Spectral analysis was systematically carried out between 250 and 8,000 Hz. Cutoff frequencies were 250, 1,676, and 8,000 Hz for the 2-band condition; 250, 709, 1,676, 3,713, and 8,000 Hz for the 4-band condition; 250, 437, 709, 1,104, 1,676, 2,507, 3,713, 5,462, and 8,000 Hz for the 8-band condition; and 250, 335, 437, 561, 709, 888, 1,104, 1,363, 1,676, 2,053, 2,507, 3,054, 3,713, 4,506, 5,462, 6,613, and 8,000 Hz for the 16-band condition. For each filtered frequency band signal, temporal envelope was extracted by half-wave rectification and envelope smoothing with a 500-Hz low-pass third-order IIR elliptical filter. The extracted temporal envelope was then used to modulate white noise delivered by a pseudorandom number generator, and the resulting signal was filtered through the same sixth-order IIR elliptical filter that was used for the frequency band selection. Finally, signals obtained from each frequency band were recombined additively, and the overall acoustic level was readjusted to match the original sound level. The performance of at least 10 subjects was analyzed for each band condition. In a last protocol, NH subjects ($n = 42$) were tested on three lists of 20 disyllabic words presented in V conditions.

In all conditions, the lists of words were equalized for syllabic structure (CV/CVC/CCV), language utilization frequency (Brulex), and anterior–posterior phonemic constitution. The stimuli were uttered by the female French speech therapist, who pronounced each word with even intonation, tempo, and vocal intensity. Utterances were recorded in an anechoic chamber with a professional digital video camera with lights focused on the face such that minimal shadowing occurred. Video was digitized at 720 × 576 pixels at 25 frames per sec, and sound was digitized at 32,000 Hz by using a 16-bit quantization. Audiovisual stimuli with sound degradation were made by using Adobe Premiere Pro 7.0 (Adobe Systems, Mountain View, CA), and temporal coincidence was respected between the original and processed sounds. All stimuli were finally exported in MPEG2 video format with maximum encoding quality.

Visual and Auditory Integration Models. An increase in multisensory performance does not necessarily prove that subjects integrate their visuoauditory inputs. Indeed, being in the presence of two signals rather than one automatically increases the probability of recognition because a word can be recognized from one or the other signal. In model 1 (minimal integration), we suppose that the subjects recognize a word from what they see or hear, with two distinct and independent modules. In this case, the word will not be recognized if and only if both the auditory and visual modules fail to recognize it. If we call P_A the probability of recognizing a word from the auditory module and P_V the probability of recognizing a word from the visual module, then the probability of failure in the presence of both visual and auditory sensory input is $(1 - P_A)(1 - P_V)$. From this latter result, we can conclude that the probability of word recognition in the presence of the two modalities without integration $\hat{P}_{VA}^{(1)}$ is given by $\hat{P}_{VA}^{(1)} = P_V + P_A - P_V P_A$. Model 1 fails to account for

the multisensory enhancement in performance in all conditions (red dotted line in Fig. 3). From this latter result, we can conclude that the visual and auditory modalities are indeed combined in a word-recognition task, albeit to a weaker extent in the case of NH subjects listening to the vocoder simulation.

To quantify multisensory integration for word recognition, we used a very simplified model where a word is "recognized" when a sufficiently large number of "cues" specific to this word is detected, either from the visual or auditory input (model 2: optimal multisensory integration). For example, a threshold of six means that six or more specific cues need to be detected to identify a word. These cues could be a specific motion of the mouth or a particular pattern in the time/frequency spectrum of the auditory signal (they do not necessarily correspond to phonemes). Moreover, we assume that the quality of the sensory input controls the average number of "cues" that can be detected in this condition.

The detection of each cue is probabilistic. We suppose that each cue is detected with a particular probability P independently of the other cues. The resulting average number of detected cues is $\lambda = NP$, where N is the total number of cues present in the word and P is the probability of detecting each of them. P depends on the quality of the sensory signal and controls the performance of the model. If N is sufficiently large and P is sufficiently small, the number of detected cues on each trial, n, follows approximately a Poisson law: The probability of detecting k cues becomes $\Pi_\lambda(n = k) = (\lambda^k e^{-\lambda})/k!$. The probability of recognition corresponds to the probability that the number of detected cues will exceed a particular fixed threshold, i.e., $\Pi_\lambda(n > T)$. Thus, it is a function of both λ and T. Assuming a fixed threshold $T = 6$ for each subject and condition, we can infer the mean number of cues λ_V, λ_A, and λ_{VA} and detect the V, A, and VA sensory inputs for each subject in each condition. For example, λ_A is the value for which $P_A = \Pi_{\lambda_A}(n > T)$ is equal to the observed auditory performance.

If visual and auditory inputs were combined optimally, they should add up together (the total number of detected cues is the sum of the visually and orally detected cues). Thus, the total signal should follow a Poisson law with mean $\hat{\lambda}_{VA} = \lambda_V + \lambda_A$. From this, we can infer the performance of an ideal observer: The bisensory recognition probability corresponds to the probability that a Poisson-distributed signal with mean $\hat{\lambda}_{VA}$ will exceed the threshold T. Thus, in the case of optimal multisensory integration, we have $\hat{P}^{(2)}_{VA} = \Pi_{\lambda_V+\lambda_A}(n > T)$.

Model 2 has one "free" parameter, the threshold T. The higher the T, the stronger the multisensory enhancement compares to unisensory performance. For $T = 1$, model 2 is equivalent to model 1 (there is no true integration: A word is detected by one or the other sensory modality). $T = 6$ was used for generating the model predictions because it is the best match for both the NH subjects in listening to noise-masked speech (Fig. 3*D*) and CI users (Fig. 3 *A* and *B*). By "best match," we mean that $T = 6$ minimized the mean squared error between the model predictions and the subject-per-subject performance in these conditions. Although the presence of a free parameter prevents us from proving that multisensory integration is optimal in an absolute sense, it provides a rigorous comparison between multisensory integration performance for different conditions and subject groups.

We thank M. L. Laborde and D. Sauvajon for help collecting the data; and C. James, S. Thorpe, and F. Polleux for comments and corrections on the manuscript. This work was supported by grants from the Action Concertée Incitative Neurosciences Intégratives et Computationnelles (to J.R., O.D., and P.B.), the Fondation pour la Recherche Médicale (to S.L.), and the Action Thématique et Incitative sur Programmes et Équipes program of the CNRS (to P.B.).

1. De Gelder B, Bertelson P (2003) *Trends Cogn Sci* 7:460–467.
2. Welch RB, Warren DH (1986) in *Handbook of Perception and Human Performance*, eds Boff KR, Kaufman L, Thomas JP (Wiley, New York), Vol 1, pp 1–36.
3. McDonald JJ, Teder-Salejarvi WA, Hillyard SA (2000) *Nature* 407:906–908.
4. Vroomen J, de Gelder B (2000) *J Exp Psychol Hum Percept Perform* 26:1583–1590.
5. Stein BE, Meredith MA (1993) *The Merging of the Senses* (MIT Press, Cambridge, MA).
6. Summerfield Q (1992) *Philos Trans R Soc Lond B Biol Sci* 335:71–78.
7. McGurk H, MacDonald J (1976) *Nature* 264:746–748.
8. Sumby WH, Pollack I (1954) *J Acoust Soc Am* 26:212–215.
9. Helfer KS, Freyman RL (2005) *J Acoust Soc Am* 117:842–849.
10. Bernstein LE, Auer ET, Takayanagia S (2004) *Speech Communication* 44:5–18.
11. Schorr EA, Fox NA, van Wassenhove V, Knudsen EI (2005) *Proc Natl Acad Sci USA* 102:18748–187450.
12. Copeland BJ, Pillsbury HC, III (2004) *Annu Rev Med* 55:157–167.
13. United Kingdom Cochlear Implant Study Group (2004) *Ear Hear* 25:310–335.
14. Rauschecker JP, Shannon RV (2002) *Science* 295:1025–1029.
15. Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) *Nature* 352:236–238.
16. Fu QJ, Shannon RV, Wang X (1998) *J Acoust Soc Am* 104:3586–3596.
17. Brozinsky CJ, Bavelier D (2004) *Brain Res Cogn Brain Res* 21:1–10.
18. Neville HJ, Lawson D (1987) *Brain Res* 405:268–283.
19. Kaiser AR, Kirk KI, Lachs L, Pisoni DB (2003) *J Speech Lang Hear Res* 46:390–404.
20. Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2006) *Cereb Cortex*, in press.
21. Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) *Science* 270:303–304.
22. Cienkowski KM, Carney AE (2002) *Ear Hear* 23:439–449.
23. Stanford TR, Quessy S, Stein BE (2005) *J Neurosci* 25:6499–6508.
24. Grant KW, Walden BE, Seitz PF (1998) *J Acoust Soc Am* 103:2677–2690.
25. Tyler RS, Parkinson AJ, Woodworth GG, Lowder MW, Gantz BJ (1997) *J Acoust Soc Am* 102:508–522.
26. Munson B, Nelson PB (2005) *J Acoust Soc Am* 118:2607–2617.
27. Bernstein LE, Auer ET, Moore JK (2004) in *The Handbook of Multisensory Processes*, eds Calvert G, Spence C, Stein BE (MIT Press, Cambridge, MA), pp 203–223.
28. Ross CA, Pearlson GD (1996) *Trends Neurosci* 19:171–176.
29. Schwartz JL, Berthommier F, Savariaux C (2004) *Cognition* 93:B69–B78.
30. Fieger A, Roder B, Teder-Salejarvi W, Hillyard SA, Neville HJ (2006) *J Cogn Neurosci* 18:149–157.
31. Gougoux F, Lepore F, Lassonde M, Voss P, Zatorre RJ, Belin P (2004) *Nature* 430:309.
32. Bavelier D, Neville HJ (2002) *Nat Rev Neurosci* 3:443–452.
33. Merabet LB, Rizzo JF, Amedi A, Somers DC, Pascual-Leone A (2005) *Nat Rev Neurosci* 6:71–77.
34. Roder B, Rosler F (2004) in *The Handbook of Multisensory Processes*, eds Calvert G, Spence C, Stein BE (MIT Press, Cambridge, MA), pp 719–747.
35. Rauschecker JP (1995) *Trends Neurosci* 18:36–43.
36. Giraud AL, Truy E, Frackowiak RS, Gregoire MC, Pujol JF, Collet L (2000) *Brain* 123:1391–1402.
37. Giraud AL, Price CJ, Graham JM, Truy E, Frackowiak RS (2001) *Neuron* 30:657–663.
38. Green KM, Julyan PJ, Hastings DL, Ramsden RT (2005) *Hear Res* 205:184–192.
39. Nishimura H, Doi K, Iwaki T, Hashikawa K, Oku N, Teratani T, Hasegawa T, Watanabe A, Nishimura T, Kubo T (2000) *Neuroreport* 11:811–815.
40. Demonet JF, Thierry G, Cardebat D (2005) *Physiol Rev* 85:49–95.
41. Lagleyre S, Rouger J, Laborde ML, Demonet JF, Fraysse B, Deguine O, Barone P (2006) in *Second Meeting of the European Societies of Neuropsychology* (Eur Soc Neuropsy, Boulogne Billancourt, France), p 162.
42. Cappe C, Barone P (2005) *Eur J Neurosci* 22:2886–2902.
43. Falchier A, Clavagnier S, Barone P, Kennedy H (2002) *J Neurosci* 22:5749–5759.
44. Schroeder CE, Foxe J (2005) *Curr Opin Neurobiol* 15:454–458.
45. Gilbert G, Lorenzi C (2006) *J Acoust Soc Am* 119:2438–2444.
46. Qin MK, Oxenham AJ (2003) *J Acoust Soc Am* 114:446–454.
47. Seitz AR, Kim R, Shams L (2006) *Curr Biol* 16:1422–1427.
48. Frassinetti F, Bolognini N, Bottari D, Bonora A, Ladavas E (2005) *J Cogn Neurosci* 17:1442–1452.
49. Grantham DW, Ashmead DH, Ricketts TA (2004) in *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*, eds Pressnitzer D, de Cheveigne A, McAdams S, Collet L (Springer, New York), pp 390–397.
50. Schoen F, Mueller J, Helms J, Nopp P (2005) *Otol Neurotol* 26:429–437.
51. Greenwood DD (1990) *J Acoust Soc Am* 87:2592–2605.

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

McGurk effects in cochlear-implanted deaf subjects

Julien Rouger[a], *Bernard Fraysse*[b], *Olivier Deguine*[a,b], *Pascal Barone*[a,*]

[a]*Centre de Recherche Cerveau et Cognition, UMR CNRS 5549, Faculté de Médecine de Rangueil, 31062 Toulouse Cedex 9, France*
[b]*Service ORL, CHU Toulouse-Purpan, 1 place du Docteur Baylac, 31059 Toulouse Cedex, France*

ARTICLE INFO

Article history:
Accepted 10 October 2007
Available online 26 October 2007

Keywords:
Cochlear implantation
Audiovisual speech
Cross-modal reorganization
McGurk effect
Articulatory feature

ABSTRACT

Cochlear implants are neuroprostheses designed to restore speech perception in case of profound bilateral hearing loss. As speech is fundamentally an audiovisual percept, a deficit in processing auditory information might lead to changes in audiovisual integration of speech comprehension. Using vowel–consonant–vowel stimuli under unimodal, audiovisual congruent and audiovisual incongruent (McGurk) conditions, we tested postlingually deaf cochlear-implanted (CI) users and normally hearing (NH) subjects in order to investigate their audiovisual perceptive strategies. Mode/Place-of-articulation perceptive analysis and information transmission analysis of congruent and incongruent percepts indicated a similar sensory specialization for CI users when compared to NH subjects, with voicing and nasality cues transmitted via audition and place cues principally transmitted via vision. NH as well as CI subjects underwent typical McGurk illusory percepts. However, while normally hearing subjects show a well-balanced bimodal integration of incongruent speech, we demonstrated that cochlear implantees present a bias toward a visual-predominant bimodal integration. Our results are complementary to previous studies showing that CI users maintain a high level of speechreading, even after several years of recovery of auditory speech comprehension. Altogether, our results suggest a cross-modal reorganization of speech comprehension in cochlear-implanted patients that might recruit more strongly than in NH the visual and visuo-auditory brain areas involved in speechreading.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

1.1. *Audiovisual speech perception*

It is now clearly established that speech perception involves both visual and auditory processes. While speech information can mostly be obtained through the auditory channel, numerous studies provide evidence that visual information could improve speech intelligibility when presented simultaneously with a noisy or degraded acoustic signal (Sumby and Pollack, 1954; Miller and Nicely, 1955; Neely, 1956; Erber, 1969, 1975; Binnie et al., 1974; Summerfield, 1979; MacLeod and Summerfield, 1987; Grant and Braida, 1991; Benoit et al., 1994; Ross et al., 2007). Moreover, visual information is useful for speech detection and intelligibility, even when acoustic information is perfectly clear. For example, visual cues decrease the acoustic level of speech detection (Grant and Seitz, 2000) and allow for a better perception of speech when presented in an unfamiliar foreign language or pronounced with a foreign accent, and when speech is perfectly clear but

* *Corresponding author*. Fax: +33 562 172 809.
E-mail address: pascal.barone@cerco.ups-tlse.fr (P. Barone).

Abbreviations: CI, cochlear implant; NHS, normally hearing subjects; IT, information transmission; A, auditory only; V, visual only; AV, audiovisual; AVc, congruent audiovisual; AVnc, incongruent audiovisual (McGurk)

0006-8993/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2007.10.049

characterized by a high level of semantic complexity (Reisberg et al., 1987; Kim and Davis, 2003).

Similarly, visual information is particularly important for speech perception in developing children. For instance, young infants are highly sensitive to spatial (Aronson and Rosenbloom, 1971; Spelke and Owsley, 1979), temporal (Dodd, 1979; Pickens et al., 1994; Spelke and Cortelyou, 1980; Walker, 1982) and phonetic congruency between auditory and visual components of speech. Consequently, it has been suggested that speech acquisition in normally hearing infants is facilitated by temporal and spatial audiovisual congruency of natural speech (Erber, 1972; Kuhl and Meltzoff, 1982, 1984, 1988; MacKain et al., 1983; Walton and Bower, 1993; Legerstee, 1990; Arnold and Kopsel, 1996; Schorr et al., 2005). This audiovisual facilitation is similarly observed in hearing-impaired infants (Erber, 1972, 1975; Arnold and Kopsel, 1996) as well as in cochlear-implanted deaf children (Staller et al., 1991; Geers and Brenner, 1994; Tyler et al., 1997; Lachs et al., 2001; Geers et al., 2003).

1.2. *The McGurk effect*

Automatic integration of visual information in audiovisual speech perception is strikingly illustrated by the McGurk effect (McGurk and MacDonald, 1976; MacDonald and McGurk, 1978). This effect demonstrates that visual speech can profoundly modify speech perception when presented simultaneously with an incongruent auditory counterpart. For example, when presenting a clear auditory /ba/ dubbed with a synchronous visual /ga/, normally hearing subjects often report hearing a /da/, even though the auditory part of the stimulus is reliably identified as /ba/ when presented in the auditory-only condition.

What is noticeable is that in most cases, the acoustic part of incongruent McGurk stimuli determines the perceived mode of articulation (e.g., voicing or nasality) whereas the visual part affects the perceived place of articulation (e.g., anterior or posterior) (McGurk and MacDonald, 1976; Brancazio, 2004). In the audiovisual integration of incongruent McGurk stimuli auditory cues for place of articulation are overwhelmed by visual cues for place of articulation. These findings can be related to unimodal auditory speech studies showing a rapid degradation of auditory identification of the place of articulation when normally hearing people are presented with a degraded speech signal (Miller and Nicely, 1955; Halle et al., 1957; Shannon et al., 1995; Boothroyd et al., 1996). In hearing-impaired and cochlear-implanted users confronted with an auditory speech signal, a similar degradation of place identification is observed, accompanied by an impairment in the identification of voicing and nasality cues (Dowell et al., 1982; Danhauer et al., 1986; Tyler et al., 1986; Rosen et al., 1989; Dorman et al., 1990; Rabinowitz et al., 1992; McKay and McDermott, 1993; Pelizzone et al., 1999; Skinner et al., 1999; van Wieringen and Wouters, 1999; Valimaa et al., 2002). Conversely, when normally hearing and hearing-impaired subjects are presented with visual speech stimuli, they demonstrate a good ability to categorize the place of articulation, but a very poor ability to identify modes of articulation (Dowell et al., 1982; Danhauer et al., 1986; Tyler et al., 1986; Rosen et al., 1989).

1.3. *Speech comprehension in cochlear-implanted deaf patients*

Cochlear implants were originally designed to restore speech perception in people with a profound bilateral hearing loss. Numerous studies have shown that multichannel cochlear implantation allows for significant progressive improvement of word or sentence perception in unimodal auditory context (Chouard et al., 1983; Wilson et al., 1991; Calmels et al., 2003; UKCISG, 2004; Rouger et al., 2007). This improvement comes with an increase of activation of auditory related areas (Naito et al., 1995; Okazawa et al., 1996; Wong et al., 1999; Giraud et al., 2001).

However, cochlear-implanted users are very sensitive to noisy environments (Fu et al., 1998; Munson and Nelson, 2005) and recent studies have shown that to overcome this sensitivity, they maintain a high level of speechreading performance even several years after implantation (Rouger et al., 2007). Such changes are related to a considerable improvement of speech comprehension in an audiovisual context compared to an auditory-only context (Tyler et al., 1997; Kaiser et al., 2003; Moody-Antonio et al., 2005), suggesting a specific integration of auditory and visual information.

This hypothesis is in agreement with specific cross-modal activations observed in cochlear-implanted patients: during auditory speech comprehension, CI users recruits additional low-level visual cortical areas when compared to normally hearing controls (Giraud et al., 2001), indicating a specific binding between visual and auditory modalities at the cortical stage.

Taken together, these results suggest that cochlear-implanted users have developed specific visuo-auditory strategies for bimodal speech comprehension. However, perceptual mechanisms underlying integration of auditory and visual cues for speech in cochlear-implanted recipients are not well understood. In order to clarify this point, we investigated audiovisual speech perception mechanisms in cochlear implantees with a McGurk paradigm: the basic feature of this experiment involved exposing normally hearing subjects and postlingually deaf cochlear-implanted users to different meaningless speech identification situations, including visual-only, auditory-only, audiovisual congruent and audiovisual incongruent (McGurk) conditions. The empirical questions under consideration were: is there a McGurk effect in deaf cochlear-implanted users? What is the nature of visual component during audiovisual speech perception in cochlear-implanted subjects?

2. Results

2.1. *Unambiguous stimuli*

2.1.1. *Correct identifications*

Correct identification of stimuli was collected for each presentation condition (see Fig. 1A). When presented in A-only conditions, stimuli were, as expected, perfectly identified by normally hearing subjects (>99%), whereas cochlear-implanted users were particularly impaired (29.5% correct). When the responses are displayed in confusion matrices

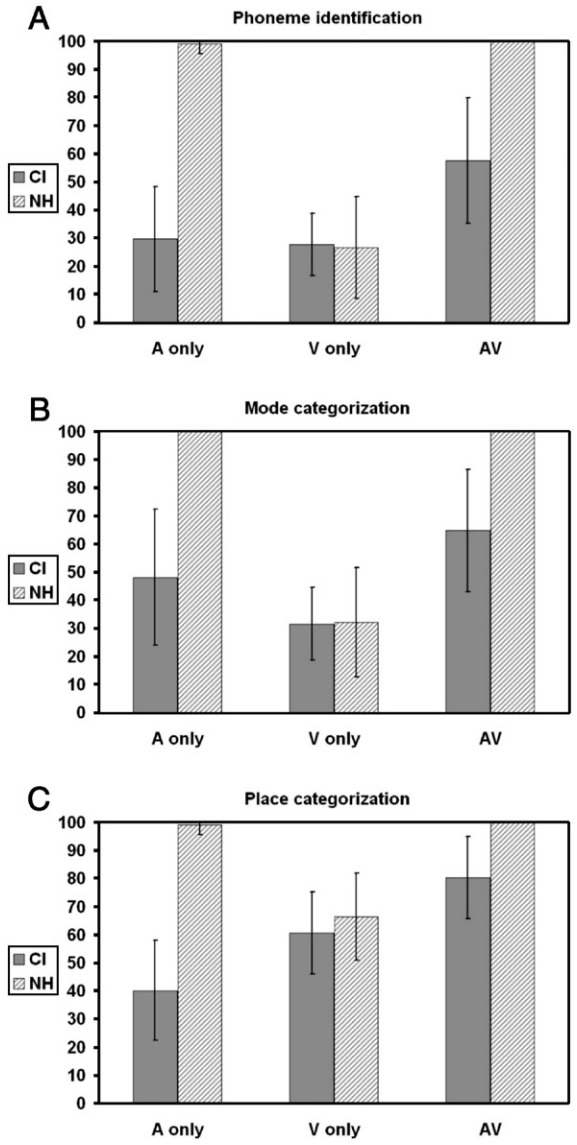


Fig. 1 – Global performance for cochlear-implanted (CI) patients and normally hearing (NH) controls in unimodal and bimodal congruent conditions. To provide valuable comparison, all histogram data are based on the set of stimuli common to A-only, V-only and AV conditions. Bars indicate standard deviations. (A) Identification scores for all subjects. (B) Global mode (i.e., manner and voicing) categorization scores for all subjects. (C) Global place categorization scores for all subjects. All three graphs show perceptive improvement in audiovisual condition when compared with unimodal conditions. Panels (B) and (C) illustrate the perceptual bias of audition and vision for, respectively, mode and place of articulation in CI patients.

(Fig. 2), NH answers are grouped on the main diagonal while CI answers only showed a weak diagonal dominance with a large spread all over the matrix, far from the diagonal.

Surprisingly, identification performance levels in the visual-only condition were identical for both populations of subjects (26.6% for NHS and 27.7% for CI users; $z=-0.85$, ns). This effect is also observed when data are presented in confusion matrices (Fig. 2) showing that the distribution of answers is grouped in a very similar way for both populations.

In the congruent audiovisual condition, stimuli were perfectly identified by the NHS group (100% correct). The CI users however had difficulty in correct identification (57.6%) although they showed significant improvement in performance when compared with the unimodal conditions ($p<0.0001$ for comparisons with A-only and V-only conditions). The congruent audiovisual confusion matrix (Fig. 3) in CI users showed a more prominent diagonal block structure than in the auditory-only condition, suggesting that visual information in the bimodal condition strongly reduced perceptive uncertainty.

2.1.2. *Global phonetic categorizations*

The non-uniform repartitions of correct identifications, especially in the visual-only condition, require further analysis to distinguish perceptual differences between the main phonetic categories. In a first approach we explored two phonetic features, the place of articulation and the mode of articulation. In this analysis, responses were sorted on the sole basis of their mode/place identity with the original stimuli (Figs. 1B and C).

For correct categorization of modes of articulation, cochlear-implanted users showed better performance levels in the audiovisual condition (64.8%), compared to other unimodal conditions (both comparisons, $p<0.0001$). Among unimodal conditions, CI users performance levels were higher in the auditory-only than in the visual-only condition (48.1% and 31.4% respectively, $p<0.0001$). The NHS group achieved perfect performance (100%) for audiovisual and auditory-only conditions, while obtaining similar scores as CI users for the visual-only condition (32.1%; $z=-0.273$, ns).

Analysis of the categorization of places of articulation showed higher performance for the CI users group in the audiovisual condition (80.3%), compared to other unimodal conditions. The comparison of performance levels in unimodal conditions revealed a better categorization of places in the visual-only condition than in the auditory-only condition (60.6% vs. 40.2%; $p<0.0001$). Perfect identification was reported in the A and AV conditions by the NH group. In the visual condition, the categorization of the place of articulation was performed at similar levels by NH and CI subjects (66.3% correct in NHS, $z=-1.224$, ns).

Categorization results in cochlear-implanted users clearly show the well-known perceptual specialization of audition and vision: modes of articulation are better identified via audition whereas places of articulation are easier to identify via vision. In the audiovisual condition, CI users obtained better results for places of articulation than for modes of articulation. This could be explained by place recognition in the V-only condition, which shows the highest performance levels among all categories (mode or place) and all unimodal presentations (compare Figs. 1B and C).

2.1.3. *Detailed phonetic categorizations: modes of articulation*

We further refined our analysis and looked for perceptual differences within the place-of-articulation and mode-of-articulation categories. In this second approach we subdivided each of the phonetic categories into three groups. Modes of articulation were decomposed into voiceless stops, voiced

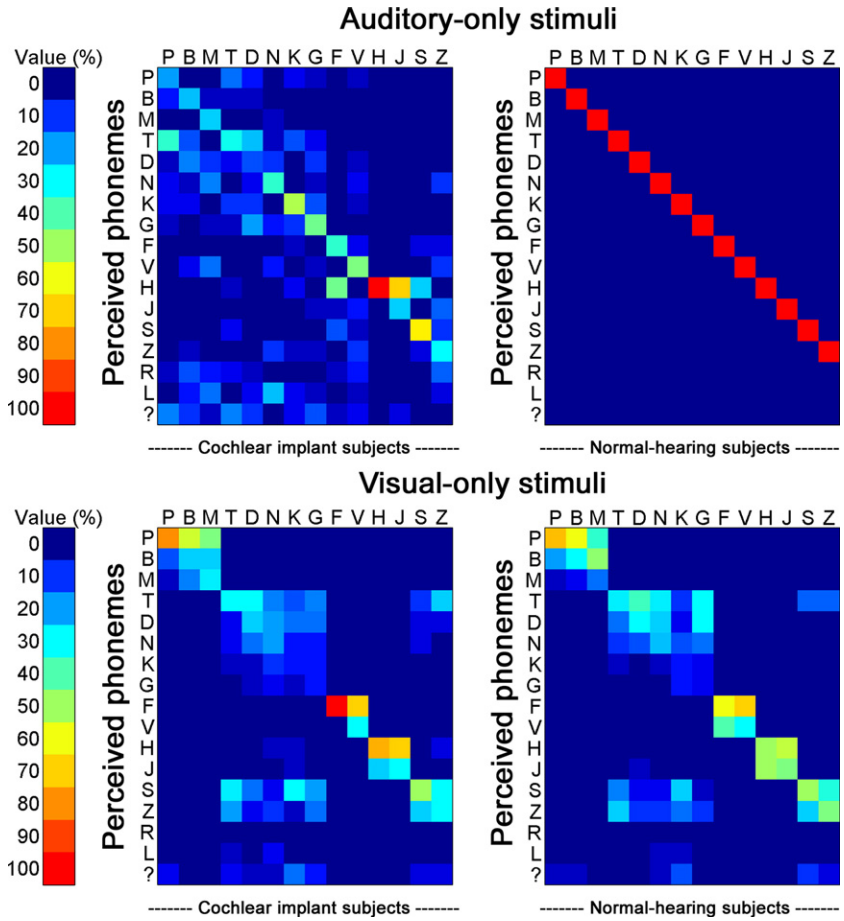


Fig. 2 – Confusion matrices for cochlear-implanted (CI) patients and normally hearing (NH) controls in auditory-only (top) and visual-only conditions (down). Stimuli are displayed on horizontal axis and answers on vertical axis; CI on left and NH on right. These confusion matrices illustrate the impaired phonetic auditory perception in CI patients (top) and the similarity of visual phonetic perception between CI and NH subjects (down).

stops and nasals, while places of articulation were decomposed into bilabials, dentals and velars. As previously, we sorted the subjects' answers on the sole basis of their parent phonetic subcategory with respect to the original stimuli.

For subcategorization of modes of articulation (Fig. 4A), CI users showed better performance in identifying voiceless stops than voiced stops or nasals, and this was true over all conditions. The last two categories were recognized at similar levels in all sensory conditions ($p>0.05$ for all comparisons). The identification of mode-of-articulation subcategories was better in the audiovisual than in the auditory-only conditions and better in the auditory-only than in the visual-only conditions over all subcategories: voiceless (AV 80.8%>A 59.6%>V 43.4%, $p<0.02$; paired test), voiced (AV 56.6%>A 42.4%>V 24.2%, $p<0.02$) and nasals (AV 53.0%>A 39.4%>V 24.2%, $p<0.04$). As mentioned previously, NH subjects presented perfect rates of recognition in the A-only and AV conditions. In V-only conditions (not shown) performance levels were equivalent to those observed in the cochlear-implanted group ($z=-0.140$, ns).

Information transmission (IT) analysis shows that mode information is very weakly transmitted by vision, for CI as well as for NH subjects: for CI users, global visual IT for mode of articulation is 6.3% (6.0% for NHS), voicing IT is 3.8% (2.7% for NHS), and nasality IT is 3.7% (1.5% for NHS). Again, few differences appear between visual-alone categorizations for cochlear-implanted and normally hearing subjects. Analysis of mode information transmission by audition shows better performance than transmission by vision for cochlear-implanted users (and, not surprisingly, perfect performance for normally hearing subjects): for CI users, global auditory IT for modes of articulation is 32.4%, voicing IT is 23.1% and nasality IT is 21.5%. Mode information transmission in the audiovisual condition shows a weak improvement when compared to that obtained in the A-only condition: global audiovisual IT for modes of articulation is 30.9%, voicing IT is 29.5% and nasality IT is 24.0%.

2.1.4. *Detailed phonetic categorizations: places of articulation*

In the A-only conditions CI users had higher performance for identification of place of articulation for velars than for bilabials (respectively 60.6% vs. 27.3%, $p<0.0004$, Fig. 4B). This hierarchy was inverted when patients were tested in visual-only conditions (bilabials 97.8%>dentals 55.6%>velars 12.1%, $p<0.0001$). Bilabials were better identified for vision than for audition ($p<0.0001$), whereas velars were better identified for

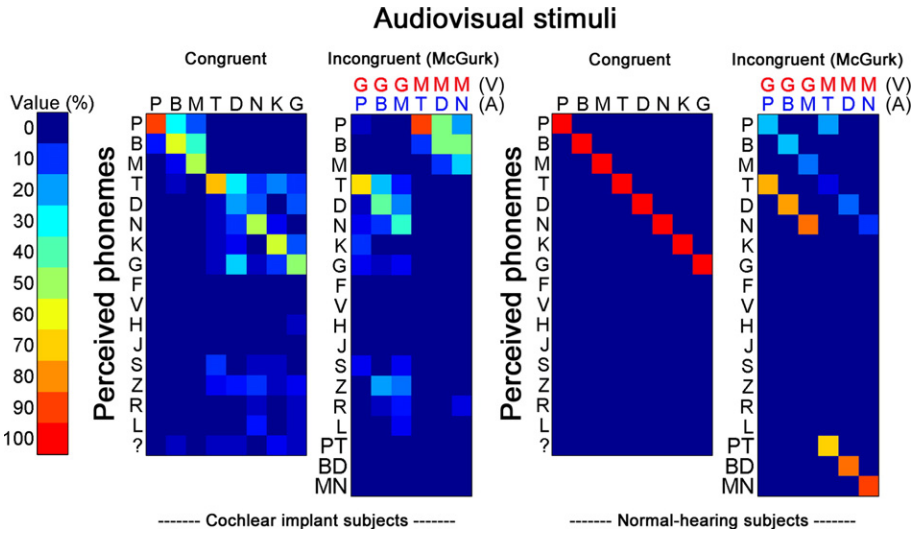


Fig. 3 – Confusion matrices for cochlear-implanted (CI) patients and normally hearing (NH) controls in audiovisual congruent and incongruent (McGurk) conditions. Stimuli are displayed on horizontal axis and answers on vertical axis; CI on left and NH on right. For incongruent stimuli, red letters refer to visual part whereas blue letters refer to auditory part. These confusion matrices illustrate the improved phonetic perception in audiovisual congruent condition and the differences in audiovisual strategies used by NH and CI subjects in incongruent condition.

audition than for vision ($p<0.0001$). As for modes of articulation, CI identification for place-of-articulation subcategories was better in the audiovisual condition compared to other unimodal conditions. Again, visual-only performance was similar for both NH and CI groups ($p>0.05$ for any comparison).

These visual-only results are well illustrated in confusion matrices (Fig. 2): for both CI users and NH subjects, a diagonal block structure is present, indicating that phonetic groups bounded in such diagonal blocks are perceptually stable for vision. Yet inside these groups the phonemes remain highly confusable. Both NH and CI subjects can easily distinguish non-labial from labial phonemes, but they can hardly refine their percept and discriminate different posterior phonemes like alveolars, dentals or velars. The visually stable groups observed are consistent with previous studies, showing high confusability between phonemes which share common place-of-articulation properties (see, e.g., Walden et al., 1977).

Information transmission analysis shows that place information is weakly transmitted by audition for cochlear-implanted subjects. Global auditory IT for places of articulation is 17.5% for CI. When detailing different categorizations, the higher score obtained for velars (21.7% vs. 5.4% for labials and 1.8% for dentals) confirms again that audition transmits information more efficiently for most posterior places of articulation. Analysis of place information transmission by vision shows better performance than transmission by audition for

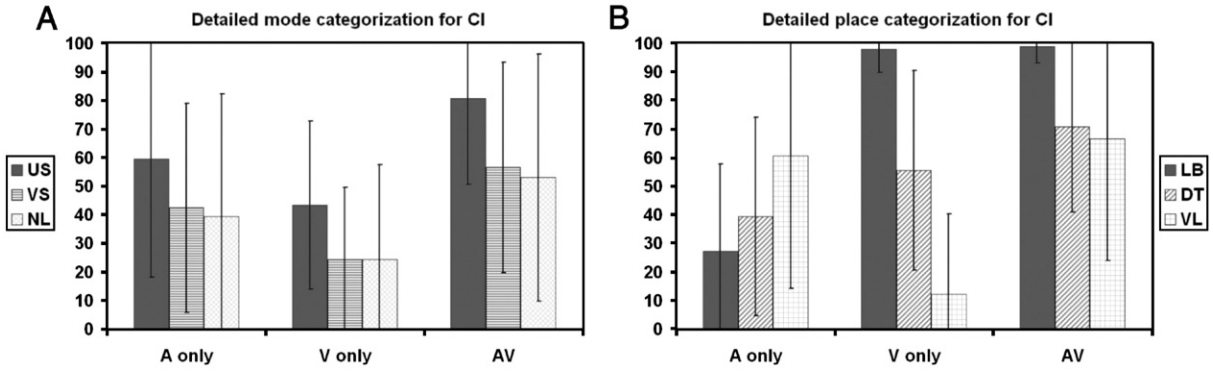


Fig. 4 – Detailed phonetic analysis for cochlear-implanted patients' answers in unimodal and bimodal congruent conditions. To provide valuable comparison, all histogram data are based on the set of stimuli common to A-only, V-only and AV conditions. Bars indicate standard deviations. (A) Detailed mode categorization for CI patients, where modes are subdivided into unvoiced stops (US), voiced stops (VS) and nasals (NL). (B) Detailed place categorization for CI patients, where places are subdivided into bilabials (LB), dentals (DT) and velars (VL). Both graphs show perceptive improvement in audiovisual condition when compared with unimodal conditions.

cochlear-implanted subjects and NHS. Global visual IT for places of articulation is 57.7% for CI users (59.0% for NH). When considering detailed places of articulation, we observed the decreasing information transmission from bilabials (96.4%) to dentals (14.9%) and velars (4.9%). It is important to notice that for cochlear-implanted subjects, information for velars is better transmitted by audition (21.7%) than by vision (4.9%), in opposition to any other place of articulation, where vision transmits information with a better efficiency than audition. Place information transmission in the audiovisual condition shows some improvement when compared to the best place IT between auditory-alone and visual-alone conditions: global visual IT for places of articulation is 66.5% in CI users.

2.2. *Ambiguous stimuli*

McGurk stimuli are purposefully ambiguous and composed of incongruent auditory and visual information. Therefore we cannot unequivocally characterize the subjects' percept and we have classified them into subcategories: perception of auditory mode, perception of visual mode (for modes of articulation), perception of auditory or visual places, fusion of auditory and visual places, combination of auditory and visual places (for places of articulation).

Normally hearing subjects typically report the percepts "apta," "abda," and "amna" for auditory dental stimuli /ata/, /ada/ and /ana/ dubbed with visual bilabial stimulus /ama/. Cochlear-implanted users respond quite differently and their typical answers were respectively "apa", "aba"/"apa" and "aba"/"ama" for the same stimuli. When the auditory bilabial stimuli /apa/, /aba/ and /ama/ are dubbed with the visual velar stimulus /aga/, both CI and NH subject groups report respectively "ata", "ada" and "ana".

2.2.1. *Perception of mode of articulation*

When the analysis is performed on perception of the mode, similar results are obtained within both experimental groups (Figs 5A and C). In normally hearing subjects, the mode of articulation (i.e., voicing and nasality) coming from the auditory part of the ambiguous stimuli is perfectly perceived (>98%), whereas the mode coming from the visual part of the stimulus is perceived at chance level (33%, which corresponds to the fact that the same visual mode is used for three different auditory modes in ambiguous stimuli). This finding is strongly

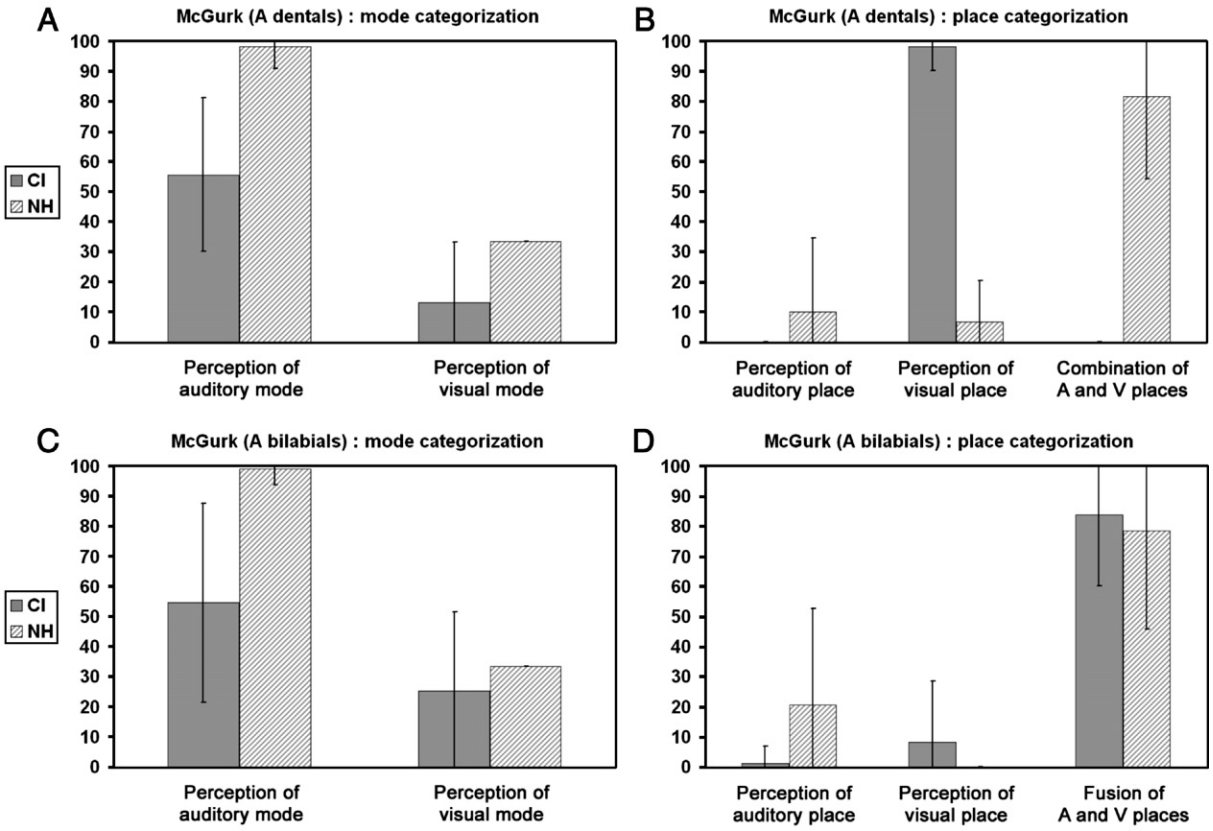


Fig. 5 – Global mode and place categorization performance for cochlear-implanted (CI) patients and normally hearing (NH) controls in audiovisual incongruent condition (McGurk). Bars indicate standard deviations. As for some incongruent stimuli, identical modes were present in both auditory and visual parts, the sum of mode categorization scores can be higher than 100%. (A) and (B) Mode and place categorization for audiovisual incongruent stimuli (McGurk) including auditory dentals /ata/, /ada/, /ana/ dubbed with a visual bilabial. (C) and (D) Mode and place categorization for audiovisual incongruent stimuli (McGurk) including auditory bilabials /apa/, /aba/, /ama/ dubbed with a visual velar. Panels (A) and (C) illustrate the impaired auditory perception of mode of articulation in CI patients. Panels (B) and (D) illustrate the differences in audiovisual strategies used by NH and CI subjects.

confirmed by information transmission analysis showing a huge difference between auditory and visual parts of the McGurk stimuli: more than 99.5% of mode information is transmitted by the auditory part of the ambiguous stimuli, whereas less than 0.1% of mode information is conveyed by the visual part of the ambiguous stimuli.

As in normally hearing subjects, cochlear-implanted users data show a higher contribution of the auditory mode compared to the visual mode. More than 54% of answers rely on the mode coming from the auditory part of the ambiguous stimuli, whereas less than 26% of them rely on the mode coming from the visual part of the ambiguous stimuli (under the 33% chance level). Again, this is confirmed by information transmission analysis, showing a transmission of more than 22% of the mode information from the auditory part, whereas the visual part conveys less than 9% of the mode information. These IT scores for mode of articulation can be compared to those obtained in auditory-only and visual-only conditions, respectively 32% and 6%.

2.2.2. *Perception of place of articulation*

For auditory dentals /ada/, /ata/, /ana/ dubbed with visual bilabial /ama/, place identification results (Fig. 5B) indicate that CI as well as NH subjects did not rely on the sole basis of the auditory part of the stimulus for perception of the place of articulation. In 82% of cases NH subjects' responses relied both on the auditory and the visual part, corresponding to hybrid percepts involving a combination of both places of articulation of the auditory and the visual part of the incongruent stimuli (e.g., bilabial–dental answer /apta/ for auditory dental /ata/ dubbed with visual bilabial /ama/). Only 10% of NH percepts were purely auditory (e.g., dental answer /ata/), whereas in 7% of them it corresponded to purely visual percepts (e.g., bilabial answer /apa/). CI users' answers were impressively driven by the visual modality: more than 98% of them reported purely visual percepts (e.g., bilabial answers /apa/ /aba/ /ama/), the place of articulation of their answers corresponding to the place of articulation of the visual part of the ambiguous stimuli.

For auditory bilabials /aba/, /apa/, /ama/ dubbed with the visual velar /aga/, place identification results (Fig. 5D) indicate again that CI as well as NH subjects did not rely on the sole basis of the auditory part of the stimulus for perception of the place of articulation. About 80% of percepts in the two groups relied both on the auditory and the visual part, responding hybrid percepts involving a fusion of places of articulation of the auditory and the visual part of the incongruent stimuli (e.g., dental answer /ata/ for auditory bilabial /apa/ dubbed with visual velar /aga/). Among non-hybrid answers, cochlear-implanted subjects relied more often on vision (CI 8% vs NH 0%) whereas normally hearing subjects rely more often on audition (NH 20.5% vs CI 1%).

As our incongruent stimuli had, by construction, a systematic correspondence of places of articulation between auditory and visual parts (i.e., auditory bilabials were systematically associated with visual velars, while auditory dentals were systematically associated with visual bilabials), there is a 100% correlation between auditory and visual places for these specific stimuli. Therefore information transmission analysis could not be used appropriately for perception of places of articulation in our McGurk stimuli, so we did not perform any IT analysis in this peculiar case.

2.2.3. *Audiovisual integration of mode and place of articulation*

We performed a correlation analysis between incongruent and congruent audiovisual percepts, comparing each AVnc answer distribution with each AVc answer distribution, using root mean square (RMS) differences. This analysis indicated that for any given incongruent (McGurk) stimulus, cochlear-implanted users' answers show the highest correlation with answers associated with the congruent audiovisual stimulus matching: 1—the mode of articulation (voicing and nasality) from the auditory part of the McGurk stimulus; 2—the place of articulation (labiality or non-labiality) from the visual part of the McGurk stimulus. Thus answers given for auditory bilabials /apa/, /aba/ and /ama/ incongruently dubbed with visual /aga/ were actually best correlated with answers given for, respectively, congruent audiovisual /ata/ (min. RMS 0.8%), /ada/ (min. RMS 5.2%) and /ana/ (min. RMS 7.3%). Answers given for auditory dentals /ata/, /ada/ and /ana/ incongruently dubbed with visual /ama/ were respectively best correlated with answers given for congruent audiovisual /apa/ (min. RMS 3.1%), /aba/ (min. RMS 8.6%) and /ama/ (min. RMS 4.6%).

This finding is visually illustrated by the confusion matrices in Fig. 3 for cochlear implantees: answers for the congruent auditory stimuli /apa/, /aba/, /ama/ (left side of the congruent CI matrix) are similarly distributed when compared to answers for the incongruent auditory stimuli /ata/, /ada/, /ana/ (right side of the incongruent CI matrix); answers for the congruent auditory stimuli /ata/, /ada/, /ana/ (middle of the congruent CI matrix) are similarly distributed when compared to answers for the incongruent auditory stimuli /apa/, /aba/, /ama/ (left side of the incongruent CI matrix).

In this experiment, non-labiality for the visual part of the ambiguous stimuli was frequently associated with dentals (>78% for both groups), and less frequently with velars (8%), although the original visual part was a velar (/aga/). This result can be linked to the weak visual identification score for velars (among NH as well as CI subjects) and their high confusibility with dentals. Further experiments have to be carried out to investigate the auditory and visual contributions to the discrimination of non-labial places in incongruent audiovisual stimuli (e.g., using only auditory and visual non-labial phonemes, like /ada/, /ata/, /ana/, /aka/, /aga/).

3. Discussion

Investigations on audiovisual speech perception mechanisms for normally hearing subjects tend to demonstrate that cross-modal interactions between vision and audition can occur very early to enhance speech intelligibility, even at a phonological level (Sams et al., 1998; Schwartz et al., 2004). Consequently, as the McGurk effect allows a discrimination of the respective contributions of auditory and visual signals in audiovisual speech perception for simple phonetic stimuli, it provides a useful tool for investigations in the field of audiovisual integration of speech.

3.1. *Uni- and bisensory integration of speech*

Our phonetical analyses of the percepts of cochlear-implanted users for unimodal stimuli are in accordance with the well-known perceptual specialization of audition and vision (Miller and Nicely, 1955). Voicing and nasality, which are typical mode-of-articulation features, convey very little perceptual information through vision but convey 5 to 6 times more information through the cochlear implant. On the other hand, labiality, which is a place-of-articulation feature, is almost perfectly identifiable with visual information while it conveys very little perceptual information through the cochlear implant.

In the visual modality, the ability to categorize the place of articulation is very accurate in assessing labiality (i.e., anterior consonants) or non-labiality (i.e., posterior consonants) but fails strongly when trying to make a discrimination among the different types of posterior consonants, like dentals and velars. Conversely, auditory information on anterior places of articulation is very poor through the cochlear implant but increases strongly for the most posterior phonemes. This can be interpreted as an inversion of audiovisual complementarity for the perception of places of articulation, where vision is helpful for audition in recognition of most anterior phonemes (bilabials), and audition is helpful for vision in recognition of most posterior phonemes (velars).

In audiovisual integration of consonants, the perceived voicing and nasality are directly derived from the auditory component of the bimodal stimulus, knowing that these articulatory features can hardly be inferred from the visual counterpart (see, e.g., Green and Kuhl, 1991; Brancazio, 2004). This mechanism of integration is similar for both normally hearing and cochlear-implanted subjects. However, in CI users, voicing and nasality can eventually be misinterpreted due to the degradation of the acoustic signal by the signal processing performed by the neuro-prosthesis. Hence the voicing and nasality of the integrated percept will rely entirely on this degraded acoustic information. This was confirmed by our correlation analysis that indicated the distribution of answers for a given incongruent audiovisual stimuli is best correlated with the distribution of answers given for the congruent audiovisual stimuli matching the voicing and nasality from the auditory part of the ambiguous stimulus (and the labiality or non-labiality from the visual part of the ambiguous stimulus).

Conversely, the perceived place of articulation is derived from both auditory and visual components of the bimodal stimulus, for normally hearing as well as for cochlear-implanted subjects. Normally hearing people perceive a place of articulation which is well balanced between auditory and visual information for place, eventually making combinations or fusions of both places. In contrast, cochlear-implanted people overweight the visual information for place and underweight the auditory information for place, showing a visual-predominant mechanism for audiovisual integration of place of articulation.

The population of CI users that participated in our protocols was very heterogeneous in age, duration of deafness and duration of implantation. Hence it was difficult to correlate any of these parameters to the percept reported by the patients during the McGurk presentation. However, we noticed a tendency in patients tested during the first 3 months after implantation to have an audiovisual percept similar to the one obtained in the visual condition, independently of their auditory scores. These preliminary results could shed some light on the time course for restoration of audiovisual integration in cochlear-implanted subjects during the first months post-implantation, in agreement with our results of audiovisual speech comprehension (Rouger et al., 2007).

3.2. *Speech perception in cochlear-implanted users*

Our results show that cochlear-implanted and normally hearing subjects present the same level of identification of stimuli during the visual-only condition. This similarity occurs at three levels of analysis: for correct identification performance as well as for global and detailed phonetic categorization, making this finding very consistent. At first glance, these results might be in contradiction with the phenomenon of intermodal compensatory mechanism that occurs following sensory deprivation (Neville, 1990; Roder et al., 1999; Bavelier and Neville, 2002; Neville and Bavelier, 2002; Bavelier et al., 2006). It is now well established that congenitally deaf patients develop some specific visual skills that affect more predominantly the spatial domain in patients using the sign language (Neville and Lawson, 1987; Proksch and Bavelier, 2002; Brozinsky and Bavelier, 2004). Similarly, postlingually deaf patients present high speech-reading capacities (Tyler et al., 1997; Kaiser et al., 2003) that remain higher than those observed in normal hearing subjects, even several years after cochlear implantation (Rouger et al., 2007). However, in the present study, stimuli were short meaningless phonetic structures, without semantic cues. In such a context, cochlear-implanted users showed the same visual phonetic abilities as normally hearing subjects. On the other hand, when speech stimuli include a semantic context (like complete words or sentences), CI users show higher visual identification levels (Rouger et al., 2007; Moody-Antonio et al., 2005; UKCISG, 2004) than normally hearing subjects. This suggests that their improved visual ability in speech recognition relies on top-down influences involving high cognitive stages of integration of the auditory stimulus, like lexical, semantic and syntaxic levels.

Because of the coarse information provided by the cochlear prosthesis (Fishman et al., 1997), correct identification of phonetic stimuli is quite low in both auditory and visuo-auditory conditions. These recognition levels remain lower than those obtained when presenting meaningful disyllabic words (see, e.g., Rouger et al., 2007), providing supplementary evidence of the benefit provided by lexical cues. In CI users, identification and categorization scores during congruent bisensory presentations are systematically higher than their best unimodal scores, regardless of whether they are performing better in vision or in audition. This increase of performance derives from the well-known perceptual improvement provided by the multisensory integration of information (Stein and Meredith, 1993), including audiovisual speech (Schwartz et al., 2004; Moody-Antonio et al., 2005; Ross et al., 2007). Again, this improvement can be noticed for three levels of analysis: global identification, global phonetic

categorization and detailed phonetic categorization. We show here that audiovisual integration can be achieved at a low perceptual level including discrimination of places and modes of articulations.

The main difference we want to point out between audiovisual integration mechanisms for normally hearing and cochlear-implanted people is the weight given respectively to the auditory and the visual modality. Analysis of McGurk percepts shows that for both groups, nasality and voicing were deduced from the auditory part, while place of articulation was integrated from auditory and visual information. In the case of CI users, mode confusion for McGurk stimuli was consistent with mode confusion for audiovisual congruent stimuli. However, detailed analysis of perceived places of articulation for incongruent audiovisual stimuli shows that cochlear-implanted users tend to overweigh visual cues compared to auditory cues, whereas normally hearing subjects tend to balance the weights of cues coming from both modalities. Thus in the case of ambiguous stimuli, and due to an uncertainty in the auditory signal, CI perceptual decisions are based on vision, their most reliable sensory channel. This trend in favor of the visual modality can be linked to the high speech-reading skills developed by CI users during their period of deafness and maintained several years after cochlear implantation, in spite of the progressive recovery of their auditory functions (Rouger et al., 2007). As CI users are highly sensitive to noise (Fu et al., 1998; Munson and Nelson, 2005), this visually orientated strategy allows them to maintain through the mechanisms of multisensory integration a high level of speech comprehension in disturbed noisy environments.

3.3. *Neural correlates for audiovisual integration of speech in cochlear-implanted users*

In CI patients, the preservation of the auditory nerve fibers allows the cochlear neuroprosthesis to reactivate the cortical network involved in sound and speech processing. However, and probably because the speech processor of the implant provides only a degraded signal that requires compensatory strategies (Shannon, 2007), CI patients present different levels of activation in auditory areas involved in semantic and/or phonological speech processing when compared to normally hearing controls (Giraud et al., 2000). Furthermore, the present results in CI users reveal a visually dominated bimodal integration in McGurk percepts, suggesting that the cortical network involved in visuo-auditory speech comprehension might differ from the one described in normally hearing subjects. In audiovisual speech, because the visual cues precede the auditory signal, visuo-auditory interaction can be observed from early stages of auditory processing (Besle et al., 2004) up to highly integrative polymodal areas such as the superior temporal gyrus (STG; see Calvert et al., 2000; Sekiyama et al., 2003; Wright et al., 2003; Möttönen et al., 2004; Kang et al., 2006). Our behavioral data suggest that this visual influence could be more prominent in CI users. Accordingly, the cortical network involved in CI users during auditory speech comprehension recruits additional low-level visual areas (Giraud et al., 2001), confirming a specific synergy between visual and auditory systems in deaf subjects. These interactions, probably mediated through direct heteromodal connections (Cappe and Barone, 2005) are probably more efficient in CI users (Lagleyre et al., 2006).

In conclusion, cross-modal binding for audiovisual speech perception requires a distributed network of cortical (Macaluso et al., 2004; Skipper et al., 2005; Saito et al., 2005) and subcortical (Champoux et al., 2006; Musacchia et al., 2006) brain regions involved in different aspects of temporal and spatial integration. Among them, the posterior parietal cortex appears to be particularly crucial when subjects are experiencing the McGurk illusion (Jones and Callan, 2003; Hamilton et al., 2006). How this parietal cortex responds in CI users remains to be explored.

4. Experimental procedure

4.1. *Participants*

33 Cochlear-implanted (CI) patients and 39 normally hearing (NH) subjects took part in this study. Participants of both groups were native French speakers with self-reported normal or corrected-to-normal vision and without any previously known language or cognitive disorders. All cochlear-implanted patients had postlingually acquired deafness and were recipients of an unilateral nucleus (cochlear) implant (CI-22 or CI-24) with from 16 to 22 functional electrodes, using a range of different sound-coding strategies (SPEAK, RCS or ACE). Cochlear implants were completely inserted in the scala tympani, allowing an optimal activation of stimulating electrodes. Post-implantation time varied from 0 to 96 months (mean 21.4±25.8 months). Performance was collected during regular visits to the ear–nose–throat department following a standard rehabilitation program. The cochlear-implanted group comprised 16 females and 17 males with a mean age of 52.0 (±15.8) years, the normally hearing group comprised 20 females and 19 males with a mean age of 32.0 (±13.0) years. All participants gave their full-informed consent prior to their inclusion in this study, in accordance with local ethics committees (Toulouse, France).

4.2. *Stimuli material*

Stimuli were meaningless vowel–consonant–vowel disyllables, where the vowel was always /a/ and the consonant was randomly chosen among the six stops /p/ /b/ /t/ /d/ /k/ /g/, the two nasals /m/ /n/ and the six fricatives /f/ /v/ /s/ /z/ /ʃ/ /3/ (/ʃ/ and /3/ will be referred to as /H/ and /J/ in figures). Six McGurk combinations were used: utterances involving the three auditory bilabials /p/, /b/ and /m/ were dubbed with the visual stimulus /aga/ (velar), while utterances involving the three auditory dentals /t/, /d/ and /n/ were dubbed with the visual stimulus /ama/ (bilabial). All stimuli were uttered by a female French speech therapist who was given the instruction to repeat each stimulus three times while maintaining a normal pronunciation with an even intonation, tempo and vocal intensity. Utterances were recorded in an soundproof chamber with a professional digital video camera. The speaker was recorded with lights focused on her face so that minimal shadowing occurs. The video recording showed the speaker facing the camera, with her entire head against a uniform yellow background. Video was digitized at 25 frames per

second with a 720×576 graphic resolution. Sound was digitized at 32,000 Hz using a 16-bit quantization.

Final stimuli were processed on a 2.4-GHz Pentium 4 processor using Adobe Premiere Pro 7.0. Each sequence of three repetitions was isolated with a short rest time before and after the utterances (mean 250 ms). Four conditions were tested: auditory-only (A), visual-only (V), congruent audiovisual (AVc) and incongruent audiovisual (AVnc, McGurk). For incongruent stimuli, audiovisual dubbing was carried out to ensure temporal coincidence between the points of burst release of each of the three utterances for the dubbing and the original auditory consonants. All stimuli were finally exported in MPEG2 video format with maximum encoding quality.

4.3. *Stimuli presentation and procedure*

CI and NH subjects were tested through two psychophysical protocols. The first one (I) used the six stops and the two nasals for auditory-only (A), visual-only (V) and audiovisual congruent (AVc) conditions, and the three bilabial McGurk stimuli for the audiovisual incongruent (AVnc) condition. The second one (II) used the six stops, the two nasals and the six fricatives for A and V conditions; the six stops and the two nasals for the AVc condition; the three bilabial and the three dental McGurk stimuli for the AVnc condition.

To minimize precedence as well as expectation biases, stimuli were randomly presented into small A, V and AV blocks (from 2 to 3 stimuli per block in protocol I and from 4 to 5 stimuli per block in protocol II) distributed in random order. McGurk stimuli were randomly included in the AV blocks, with one or none McGurk stimulus per block for protocol I and two McGurk stimuli per AV block for protocol II.

The stimuli were presented on a laptop computer with a 16-in. monitor. The video image was 8.1 in. high×10.9 in. wide and was surrounded by a black border. The sound was presented at a fixed comfortable listening level (65 dB) through good-quality loudspeakers placed on each side of the computer monitor.

Participants were tested individually. They were seated in front of the computer at a normal viewing distance (about 25 in.), thus creating a total viewing angle of approximately 30°. The participants were instructed to watch the monitor and to listen to the loudspeakers during stimulus presentation, then to repeat what they had heard. The experimenter sat in the room to check that the participants watched the monitor and to record verbatim their verbal answers.

4.4. *Data and statistical analysis*

Incongruent audiovisual percepts can be split into four categories. First, auditory-dominant percept occurs when the perceived consonant has a similar place of articulation as the auditory signal [auditory stimulus: /aba/ (bilabial); visual stimulus: /aga/ (velar); percept: /aba/ (bilabial)]. Second, visual-dominant percept occurs when the reported percept has the same place of articulation as the visual signal [auditory stimulus: /aba/ (bilabial); visual stimulus: /aga/ (velar); percept: /aga/ (velar)]. Third, combination percept corresponds to percepts involving two distinct places of articulation corresponding to both visual and auditory stimuli [auditory stimulus: /aga/ (velar); visual stimulus: /aba/ (bilabial); percept: /abga/ (bilabial + velar)]. Lastly, we classified as fusion percept when the perceived place of articulation is intermediate between both visual and auditory places of articulation [auditory stimulus: /aba/ (bilabial); visual stimulus: /aga/ (velar); percept: /ada/ (dental)]. A systematic phonetic transcription was carried out for the recorded answers. Answers without any VCV or VCCV pattern were excluded from the final analysis, which focused only on the central consonants of the reported percept. Consonants were then classified by mode and by place of articulation.

When participants happened to report a double-consonant nucleus (this occurs especially for incongruent audiovisual stimuli, e.g., /bd/), the following rule was applied:

- modes of articulation of double nucleus with two identical modes of articulation (e.g., /pt/, /bd/, /mn/) were identified to the common mode of both consonants of this nucleus (e.g., /pt/ voiceless stop, /bd/ voiced stop, /mn/ nasal);
- any other combinations were considered as ambiguous and were not included for place and mode analysis. Less than 1.3% of all audiovisual incongruent percepts happened to be ambiguous, for CI as well as for NH. All participants gave spontaneously an answer for all audiovisual incongruent stimuli.

As the null hypothesis of normality was rejected (Lilliefors and Kolmogorov–Smirnov) at the significance level of 5% for the distribution of correct categorizations in most sub-groups, non-parametric statistical analysis was carried out. The Mann–Whitney non-paired *U*-test was used to compare CI and NH performance for any common experimental condition. The Wilcoxon paired test was used to compare two different experimental conditions for a common population (CI or NH). As the results from stimuli common to both protocols were strongly similar without any statistical difference, all data were merged together.

To provide additional information on subject's perceptive mechanisms relatively to each specific articulatory feature, an information transmission (IT) analysis was performed on categorizations of places and modes of articulation, according to methods originally described by Miller and Nicely (1955). Such an analysis involves mathematical computations of the amount of information (similar to Shannon's entropy) embedded in stimuli and answers as well as in correlation between stimuli and answers. Through these computations, IT analysis allows a quantification of the amount of perceptual information that was preserved during the perception process. As few stimuli were presented to each participant, we computed IT scores over all subjects in each group for each articulatory feature in order to avoid biases of maximum likelihood estimates for frequencies (see below); therefore no individual IT scores could be calculated. Hence this IT analysis only provides general trends that call for further investigations with a large number of stimuli.

Here we recall the basic principles of IT analysis. If we call x_i the phoneme used for stimulation and f_i its presentation frequency, where i can assume k discrete values in [1 … k], then the amount of information carried through x is defined by the theory of information as the mean logarithmic probability

of x : $\mathrm{MLP}(x) = E(-\log f) = -\sum_i f_i \log f_i$. Now we call y_j the phoneme given in answer (associated with answer frequency f_j), where j can assume m discrete values in [1 … m], and we call c_{ij} the couple formed by answer y_j occurring with presentation of stimulus x_i (associated with co-occurrence frequency f_{ij}). Then a measure of covariance of stimulus x with answer y is given by: $\mathrm{cov}(x,y) = \mathrm{MLP}(x) + \mathrm{MLP}(y) - \mathrm{MLP}(c) = -\sum_{ij} f_{ij} \log \frac{f_i f_j}{f_{ij}}$. The ratio of information transmitted from stimuli x to percepts y is finally given by IT(x, y)=100*cov(x, y)/MLP(x). IT varies from 0 to 100%; 0 indicates that responses are not correlated to stimuli (random answers or same answer for every stimuli) and 100% indicates that responses are perfectly correlated to stimuli (stimuli-specific answers).

Acknowledgments

We thank the cochlear-implanted and normally hearing people for their participation to this study, Marie-Laurence Laborde and Sébastien Lagleyre for their help in stimuli recording, Jean-François Démonet and Daniel Pressnitzer for their helpful comments and Catherine Marlot and Leila Reddy for corrections on the manuscript. This work was supported by an Action Concertée Incitative grant from the French Research Ministery (Neurosciences intégratives et computationnelles, NIC 0085, june 2003) and the Fondation pour la Recherche Médicale (JR).

REFERENCES

Arnold, P., Kopsel, A., 1996. Lipreading, reading and memory of hearing and hearing-impaired children. Scand. Audiol. 25, 13–20.

Aronson, E., Rosenbloom, S., 1971. Space perception in early infancy: perception within a common auditory-visual space. Science 172, 1161–1163.

Bavelier, D., Neville, H.J., 2002. Cross-modal plasticity: where and how? Nat. Rev., Neurosci. 3, 443–452.

Bavelier, D., Dye, M.W., Hauser, P.C., 2006. Do deaf individuals see better? Trends Cogn. Sci. 10, 512–518.

Benoit, C., Mohamadi, T., Kandel, S., 1994. Effects of phonetic context on audio-visual intelligibility of French. J. Speech Hear. Res. 37, 1195–1203.

Besle, J., Fort, A., Delpuech, C., Giard, M.H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. Eur. J. Neurosci. 20, 2225–2234.

Binnie, C.A., Montgomery, A.A., Jackson, P.L., 1974. Auditory and visual contributions to the perception of consonants. J. Speech Hear. Res. 17, 619–630.

Boothroyd, A., Mulhearn, B., Gong, J., Ostroff, J., 1996. Effects of spectral smearing on phoneme and word recognition. J. Acoust. Soc. Am. 100, 1807–1818.

Brancazio, L., 2004. Lexical influences in audiovisual speech perception. J. Exp. Psychol. Hum. Percept. Perform. 30, 445–463.

Brozinsky, C.J., Bavelier, D., 2004. Motion velocity thresholds in deaf signers: changes in lateralization but not in overall sensitivity. Brain Res. Cogn. Brain Res. 21, 1–10.

Calmels, M.N., Saliba, I., Wanna, G., Cochard, N., Fillaux, J., Deguine, O., Fraysse, B., 2003. Speech perception and speech intelligibility in children after cochlear implantation. Int. J. Pediatr. Otorhinolaryngol. 68, 347–351.

Calvert, G.A., Campbell, R., Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr. Biol. 10, 649–657.

Cappe, C., Barone, P., 2005. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur. J. Neurosci. 22, 2886–2902.

Champoux, F., Tremblay, C., Mercier, C., Lassonde, M., Lepore, F., Gagne, J.P., Theoret, H., 2006. A role for the inferior colliculus in multisensory speech integration. NeuroReport 17, 1607–1610.

Chouard, C.H., Fugain, C., Meyer, B., Lacombe, H., 1983. Long-term results of the multichannel cochlear implant. Ann. N.Y. Acad. Sci. 405, 387–411.

Danhauer, J.L., Erratt, J.D., Edgerton, B.J., 1986. Auditory/Visual speech perception by cochlear implant users and hearing aid wearers. Am. J. Otol. 7, 354–360.

Dodd, B., 1979. Lipreading in infants: attention to speech presented in and out of synchrony. Cogn. Psychol. 11, 478–484.

Dorman, M.F., Soli, S., Dankowski, K., Smith, L.M., McCandless, G., Parkin, J., 1990. Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant. J. Acoust. Soc. Am. 88, 2074–2079.

Dowell, R.C., Martin, L.F., Tong, Y.C., Clark, G.M., Seligman, P.M., Patrick, J.F., 1982. A 12-consonant study on a multiple-channel cochlear implant patient. J. Speech Hear. Res. 25, 509–516.

Erber, N.P., 1969. Interaction of audition and vision in the recognition of oral speech stimuli. J. Speech Hear. Res. 12, 423–425.

Erber, N.P., 1972. Auditory, visual and auditory-visual recognition of consonants by children with normal and impaired hearing. J. Speech Hear. Res. 15, 413–422.

Erber, N.P., 1975. Auditory-visual perception of speech. J. Speech Hear. Disord. 33, 771–778.

Fishman, K.E., Shannon, R.V., Slattery, W.H., 1997. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. J. Speech Lang. Hear. Res. 40, 1201–1215.

Fu, Q.J., Shannon, R.V., Wang, X., 1998. Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. J. Acoust. Soc. Am. 104, 3586–3596.

Geers, A., Brenner, C., 1994. Speech perception results: audition and lipreading enhancement. Volta Rev. 96, 97–108.

Geers, A., Brenner, C., Davidson, L., 2003. Factors associated with development of speech perception skills in children implanted by age five. Ear Hear. 24, 24S–35S.

Giraud, A.L., Truy, E., Frackowiak, R.S., Gregoire, M.C., Pujol, J.F., Collet, L., 2000. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. Brain 123, 1391–1402.

Giraud, A.L., Price, C.J., Graham, J.M., Truy, E., Frackowiak, R.S., 2001. Cross-modal plasticity underpins language recovery after cochlear implantation. Neuron 30, 657–663.

Grant, K.W., Braida, L.D., 1991. Evaluating the articulation index for audiovisual input. J. Acoust. Soc. Am. 89, 2952–2960.

Grant, K.W., Seitz, P., 2000. The use of visible speech cues for improving auditory detection of spoken sentences. J. Acoust. Soc. Am. 108, 1197–1208.

Green, K.P., Kuhl, P.K., 1991. Integral processing of visual place and auditory voicing information during phonetic perception. J. Exp. Psychol. Hum. Percept. Perform. 17, 278–288.

Halle, M., Hughes, G.W., Radley, J.P.A., 1957. Acoustic properties of stop consonants. J. Acoust. Soc. Am. 20, 107–116.

Hamilton, R.H., Shenton, J.T., Coslett, H.B., 2006. An acquired deficit of audiovisual speech processing. Brain Lang. 98, 66–73.

Jones, J.A., Callan, D.E., 2003. Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. NeuroReport 14, 1129–1133.

Kaiser, A.R., Kirk, K.I., Lachs, L., Pisoni, D.B., 2003. Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. J. Speech Lang. Hear. Res. 46, 390–404.

Kang, E., Lee, D.S., Kang, H., Hwang, C.H., Oh, S.H., Kim, C.S., Chung, J.K., Lee, M.C., 2006. The neural correlates of cross-modal interaction in speech perception during a semantic decision task on sentences: a PET study. NeuroImage 32, 423–431.

Kim, J., Davis, C., 2003. Hearing foreign voices: does knowing what is said affect masked visual speech detection? Perception 32, 111–120.

Kuhl, P.K., Meltzoff, A.N., 1982. The bimodal development of speech in infancy. Science 218, 1138–1141.

Kuhl, P.K., Meltzoff, A.N., 1984. The intermodal representation of speech in infants. Infant Behav. Dev. 7, 361–381.

Kuhl, P.K., Meltzoff, A.N., 1988. Speech as an intermodal object of perception. In: Yonas, A. (Ed.), Perceptual development in infancy: The Minnesota Symposia on child psychology, vol. 20. Erlbaum, Hillsdale, NJ.

Lachs, L., Pisoni, D.B., Kirk, K.I., 2001. Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: a first report. Ear Hear. 22, 236–251.

Lagleyre, S., Rouger, J., Laborde, M.L., Demonet, J.F., Fraysse, B., Deguine, O., Barone, P., 2006. Role of visuo-auditory integration in speech comprehension in deaf subjects with cochlear implants. 2nd Meeting of the European Societies of Neuropsychology, p. 162.

Legerstee, M., 1990. Infants use multimodal information to imitate speech sounds. Infant Behav. Dev. 17, 829–840.

Macaluso, E., George, N., Dolan, R., Spence, C., Driver, J., 2004. Spatial and temporal factors during processing of audiovisual speech: a PET study. NeuroImage 21, 725–732.

MacDonald, J., McGurk, H., 1978. Visual influences on speech perception processes. Percept. Psychophys. 24, 253–257.

MacKain, K., Studdert-Kennedy, M., Spieker, S., Stern, D., 1983. Infant intermodal speech perception is a left hemisphere function. Science 219, 1347–1349.

MacLeod, A., Summerfield, Q., 1987. Quantifying the contribution of vision to speech perception in noise. Br. J. Audiol. 21, 131–141.

McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. Nature 264, 746–748.

McKay, C.M., McDermott, H.J., 1993. Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSP) and the Mini Speech Processor (MSP). Ear Hear. 14, 350–367.

Miller, G.A., Nicely, P., 1955. An analysis of perceptual confusions among some English consonants. J. Acoust. Soc. Am. 27, 338–352.

Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer Jr., E.T., Fisher, L., Bernstein, L.E., 2005. Improved speech perception in adult congenitally deafened cochlear implant recipients. Otol. Neurotol. 26, 649–654.

Möttönen, R., Schürmann, M., Sams, M., 2004. Time course of multisensory interactions during audiovisual speech perception in humans: a magnetoencephalographic study. Neurosci. Lett. 363, 112–115.

Munson, B., Nelson, P.B., 2005. Phonetic identification in quiet and in noise by listeners with cochlear implants. J. Acoust. Soc. Am. 118, 2607–2617.

Musacchia, G., Sams, M., Nicol, T., Kraus, N., 2006. Seeing speech affects acoustic information processing in the human brainstem. Exp. Brain Res. 168, 1–10.

Naito, Y., Okazawa, H., Honjo, I., Hirano, S., Takahashi, H., Shiomi, Y., Hoji, W., Kawano, M., Ishizu, K., Yonekura, Y., 1995. Cortical activation with sound stimulation in cochlear implant users demonstrated by positron emission tomography. Cogn. Brain Res. 2, 207–214.

Neville, H.J., 1990. Intermodal competition and compensation in development. Evidence from studies of the visual system in congenitally deaf adults. Ann. N.Y. Acad. Sci. 608, 71–91.

Neville, H., Bavelier, D., 2002. Human brain plasticity: evidence from sensory deprivation and altered language experience. Prog. Brain Res. 138, 177–188.

Neville, H.J., Lawson, D., 1987. Attention to central and peripheral visual space in a movement detection task: an event-related potential and behavioral study: II. Congenitally deaf adults. Brain Res. 405, 268–283.

Neely, K.K., 1956. Effect of visual factors on the intelligibility of speech. J. Acoust. Soc. Am. 28, 1275–1277.

Okazawa, H., Naito, Y., Yonekura, Y., Sadato, N., Hirano, S., Nishizawa, S., Magata, Y., Ishizu, K., Tamaki, N., Honjo, I., Konishi, J., 1996. Cochlear implant efficiency in pre- and postlingually deaf subjects. A study with H_2(15)O and PET. Brain 119, 1297–1306.

Pelizzone, M., Cosendai, G., Tinembart, J., 1999. Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for Ineraid implanted subjects. Ear Hear. 20, 228–237.

Pickens, J., Field, T., Nawrocki, T., Martinez, A., Soutullo, D., Gonzalez, J., 1994. Full-term and preterm infants' perception of face–voice synchrony. Infant Behav. Dev. 17, 447–455.

Proksch, J., Bavelier, D., 2002. Changes in the spatial distribution of visual attention after early deafness. J. Cogn. Neurosci. 14, 687–701.

Rabinowitz, W.M., Eddington, D.K., Delhorne, L.A., Cuneo, P.A., 1992. Relations among different measures of speech reception in subjects using a cochlear implant. J. Acoust. Soc. Am. 92, 1869–1881.

Reisberg, D., McLean, J., Goldfield, A., 1987. Easy to hear but hard to understand. In: Dodd, B., Campbell, R. (Eds.), Hearing by Eye: The Psychology of Lipreading. Erlbaum, Hillsdale, NJ, pp. 97–114.

Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S.A., Neville, H.J., 1999. Improved auditory spatial tuning in blind humans. Nature 400, 162–166.

Rosen, S., Walliker, J., Brimacombe, J.A., Edgerton, B.J., 1989. Prosodic and segmental aspects of speech perception with the House/3M single-channel implant. J. Speech Hear. Res. 32, 93–111.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., Foxe, J.J., 2007. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb. Cortex 17, 1147–1153.

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P., 2007. Evidence that cochlear implanted deaf patients are better multisensory integrators. Proc. Natl. Acad. Sci. U. S. A. 104, 7295–7300.

Saito, D.N., Yoshimura, K., Kochiyama, T., Okada, T., Honda, M., Sadato, N., 2005. Cross-modal binding and activated attentional networks during audio-visual speech integration: a functional MRI study. Cereb. Cortex 15, 1750–1760.

Sams, M., Manninen, P., Surakka, V., Helin, P., Kättö, R., 1998. McGurk effect in Finnish syllables, isolated words, and words in sentences: effects of word meaning and sentence context. Speech Commun. 26, 75–87.

Schorr, E.A., Fox, N.A., van Wassenhove, V., Knudsen, E.L., 2005. Auditory-visual fusion in speech perception in children with cochlear implants. Proc. Natl. Acad. Sci. U. S. A. 102, 18748–18750.

Schwartz, J.L., Berthommier, F., Savariaux, C., 2004. Seeing to hear better: evidence for early audio-visual interactions in speech identification. Cognition 93, B69–B78.

Sekiyama, K., Kanno, I., Miura, S., Sugita, Y., 2003. Audio-visual speech perception examined by fMRI and PET. Neurosci. Res. 47, 277–287.

Shannon, R.V., 2007. Understanding hearing through deafness. Proc. Natl. Acad. Sci. U. S. A. 104, 6883–6884.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonsky, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270, 303–304.
Skinner, M.W., Fourakis, M.S., Holden, T.A., Holden, L.K., Demorest, M.E., 1999. Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies: II. Consonants. Ear Hear. 20, 443–460.
Skipper, J.I., Nusbaum, H.C., Small, S.L., 2005. Listening to talking faces: motor cortical activation during speech perception. NeuroImage 25, 76–89.
Spelke, E.S., Cortelyou, A., 1980. Perceptual aspects of social knowing: looking and listening in Infancy. In: Lamb, E., Sherrod, L. (Eds.), Infant social cognition. Erlbaum, Hillsdale, NJ.
Spelke, E.S., Owsley, C., 1979. Intermodal exploration and knowledge in infancy. Infant Behav. Dev. 2, 13–27.
Staller, S.J., Dowell, R.C., Beiter, A.L., Brimacombe, J.A., 1991. Perceptual abilities of children with the nucleus 22-channel cochlear implant. Ear Hear. 12, 34S–47S.
Stein, B.E., Meredith, M.A., 1993. The Merging of the Senses. MIT Press, Cambridge, MA.
Sumby, W.H., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. 26, 212–215.
Summerfield, Q., 1979. Use of visual information for phonetic perception. Phonetica 36, 314–331.
Tyler, R.S., Preece, J.P., Lansing, C.R., Otto, S.R., Gantz, B.J., 1986. Previous experience as a confounding factor in comparing cochlear-implant processing schemes. J. Speech Hear. Res. 29, 282–287.
Tyler, R.S., Fryauf-Bertschy, H., Kelsay, D.M.R., Gantz, B.J., Woodworth, G.P., Parkinson, A., 1997. Speech perception by prelingually deaf children using cochlear implants. Otolaryngol. Head Neck Surg. 117, 180–187.
United Kingdom Cochlear Implant Study Group, 2004. Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults I: theory and measures of effectiveness. Ear Hear. 25, 310–335.
Valimaa, T.T., Maatta, T.K., Lopponen, H.J., Sorri, M.J., 2002. Phoneme recognition and confusions with multichannel cochlear implants: consonants. J. Speech Lang. Hear. Res. 45, 1055–1069.
van Wieringen, A., Wouters, J., 1999. Natural vowel and consonant recognition by Laura cochlear implantees. Ear Hear. 20, 89–103.
Walden, B.E., Prosek, R.A., Montgomery, A.A., Scherr, C.K., Jones, C. J., 1977. Effects of training on the visual recognition of consonants. J. Speech Hear. Res. 20, 130–145.
Walker, A.S., 1982. Intermodal perception of expression behaviors by human infants. J. Exp. Child Psychol. 33, 514–535.
Walton, G.E., Bower, T.G.R., 1993. Amodal representation of speech in infants. Infant Behav. Dev. 16, 233–243.
Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D. K., Rabinowitz, W.M., 1991. Better speech recognition with cochlear implants. Nature 352, 236–238.
Wong, D., Miyamoto, R.T., Pisoni, D.B., Sehgal, M., Hutchins, G.D., 1999. PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. Hear. Res. 132, 34–42.
Wright, T.M., Pelphrey, K.A., Allison, T., McKeown, M.J., McCarthy, G., 2003. Polysensory interactions along lateral regions evoked by audiovisual speech. Cereb. Cortex 13, 1034–1043.

Author's personal copy

Neuropsychologia 47 (2009) 972–979

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Improvement in speech-reading ability by auditory training: Evidence from gender differences in normally hearing, deaf and cochlear implanted subjects

K. Strelnikov[a,b], J. Rouger[a,b], S. Lagleyre[c], B. Fraysse[c], O. Deguine[a,b,c], P. Barone[a,b,*]

[a] Université Toulouse, CerCo, Université Paul Sabatier, France
[b] CNRS, UMR 5549, Faculté de Médecine de Rangueil, Toulouse, France
[c] Service d'Oto-Rhino-Laryngologie, Hopital Purpan, Toulouse, France

ARTICLE INFO

Article history:
Received 3 March 2008
Received in revised form 20 October 2008
Accepted 23 October 2008
Available online 30 October 2008

Keywords:
Gender differences
Neural plasticity
Speech-reading
Deafness
Cochlear implantation

ABSTRACT

Several neuropsychological and neuroimaging studies on gender differences in speech processing lead to the suggestion that women use the neural network of predictive and integrative analysis of speech to a larger extent than men. During speech-reading there is indeed a lack of reliable clues for word recognition which should emphasize predictive and integrative strategies of the brain. Our study aimed to explore gender differences in deaf and cochlear implanted (CI) patients at different levels during speech-reading, for words or phonemes, that we consider, correspond to increased involvement of predictive and integrative analysis. We collected speech-reading scores in a control group of normally hearing subjects ($n = 42$) and in a group of deaf patients – who are good speech-readers – tested before, early after and late after cochlear implantation ($n = 97$). Patient groups were almost equally distributed between follow-up and new patients. In normally hearing controls, women speech-read words better than men. This difference was also observed in all patients but not in experienced cochlear implant users. We did not observe a gender difference during speech-reading of isolated phonemes neither for controls nor for patients. We conclude that the better speech-reading ability of women for words but not for phonemes is in line with their greater use of predictive and integrative strategies for speech processing. Furthermore, we observed a progressive cross-modal compensation in male CI users after cochlear implantation which suggests a synergetic perceptual facilitation involving the visual and the recovering auditory modalities. This could lead to an improved performance in both auditory and visual modalities, the latter being constantly recruited to complement the crude information provided by the implant. Altogether, our data provide insights into cross-modal compensation in the adult brain following sensory privation.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Several studies using different experimental paradigms and methods have reported gender differences in speech processing. In brain imaging studies (Jaeger et al., 1998; Kansaku & Kitazawa, 2001; Kansaku, Yamaura, & Kitazawa, 2000; Phillips, Lowe, Lurito, Dzemidzic, & Mathews, 2001; Rossell, Bullmore, Williams, & David, 2002; Shaywitz et al., 1995), it has been shown that activations in men are usually more left-lateralized while activations in women are more diffuse and more symmetrical, although the exact patterns vary due to differences in experimental paradigms. A recent meta-analysis (Sommer, Aleman, Bouma, & Kahn, 2004) claimed no consistent gender difference in lateralization for speech processing concluding that such difference may be observed only in some, not yet defined, language tasks. However, this conclusion remains subject of debate as gender difference is observed for continuous speech processing (Kitazawa & Kansaku, 2005) and it is better viewed with conjunction analysis and in short time-windows (Ortigue, Thut, Landis, & Michel, 2005). According to our activation-verification model of speech processing (Strelnikov, 2008), this gender difference in the patterns of activations might suggest that women use more integrative and predictive cognitive strategies than men. However, to the best of our knowledge, there is no convincing evidence for a gender difference in speech comprehension performance. This therefore implies that men and women use different cognitive strategies that normally result in similar levels of speech comprehension. It is obvious that phoneme and word recognition within a certain context (e.g., in a phrase or a text) is influenced by this context. Thus, the use of predictive and integrative strategies may also depend on the linguistic context.

* Corresponding author at: CNRS-Université Paul Sabatier Toulouse 3, Centre de Recherche Cerveau et Cognition UMR 5549, Faculté de Médecine de Rangueil, 31062 Toulouse Cedex 9, France.
E-mail address: pascal.barone@cerco.ups-tlse.fr (P. Barone).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.10.017

Author's personal copy

The variety of contexts across the different studies may explain some fluctuations in the reported brain activation patterns, which is rather common in neuroimaging studies.

An important question relates to the audiovisual integration of speech, specifically, whether an integrative and predictive cognitive strategy in women could also lead to superior performance in audiovisual speech comprehension. Though there is no direct evidence for this, some studies have indicated that under certain conditions, women are more influenced by visual information than men, for example during the perception of the McGurk incongruent auditory and visual speech stimuli (Aloufy, Lapidot, & Myslobodskym, 1996; Irwin, Whalen, & Fowler, 2006). It is conceivable that the lack of reliable cues while speech-reading may emphasize the integrative and predictive cognitive strategy of women leading to superior recognition scores. In support of this, some studies have found that women are significantly better speech-readers than men (Johnson & Leslie, 1988; Watson, Qiu, Chamberlain, & Li, 1996) although some others did not find such a gender effect (Irwin et al., 2006). These differences in findings may depend on the experimental paradigm and the type of stimuli used (e.g., phonemes, words, sentences). In link with this, Jaeger et al. (1998) hypothesized that gender differences in language may emerge as the task demands increase.

In humans and animals, it is shown that the loss of a sensory modality leads to compensatory mechanisms which increase the performance of the remaining modalities (Bavelier & Neville, 2002; Bavelier, Dye, & Hauser, 2006; Putzar, Goerendt, Lange, Rosler, & Roder, 2007; Rauschecker, 1991; Roder et al., 1999). While in blind patients, cross-modal compensation can improve sensory thresholds in the remaining modalities (Goldreich & Kanics, 2003; Gougoux et al., 2004), enhanced performances in the preserved sense involve mainly attentional mechanisms (Collignon, Renier, Bruyer, Tranduy, & Veraart, 2006) especially in cases of deafness (Bavelier et al., 2000; Roder et al., 1999). In deaf people, such cross-modal compensation can be seen in an increased capacity to understand speech through the visual channel during speech-reading (Bernstein, Auer, & Tucker, 2001; Rouger et al., 2007). In deaf patients, the word comprehension performance in speech-reading is usually about 35%, while it barely exceeds 10% in healthy subjects (Rouger et al., 2007). Such a visually based compensation mechanism is accompanied by significant reorganization of the neural network involved in speech comprehension in deaf patients (Fine, Finney, Boynton, & Dobkins, 2005; Finney, Fine, & Dobkins, 2001; Lee, Truy, Mamou, Sappey-Marinier, & Giraud, 2007; Nishimura et al., 1999) as well as in cochlear implant patients (Giraud, Price, Graham, Truy, & Frackowiak, 2001; Green, Julyan, Hastings, & Ramsden, 2005; Ito et al., 2004). Interestingly, as we have shown in our previous study (Rouger et al., 2007), profoundly deaf patients who receive a cochlear implant (CI) maintain a high level of speech-reading capacity even after the recovery of speech comprehension through the auditory pathway. Due to these high levels of speech-reading performance, deaf and cochlear implanted patients enable us to explore gender differences in speech-reading following compensatory cortical reorganization.

We hypothesize that better predictive and integrative strategies in women can help them better adapt to the lacking or deteriorated auditory input. A further question is whether such strategies could also improve men's performance under some appropriate specific conditions.

To manipulate the level of predictability and thus the intensity of integrative predictions, we compared the performance in two lip-reading conditions, involving either isolated phoneme or disyllabic word recognition. The underlying hypothesis is that a phoneme embedded into a disyllabic word may generate predictions about the preceding or following phonemes, according to the word representation in linguistic memory, leading to a higher recognition performance. Such context-based linguistic predictions are expected to be much smaller in case of isolated phonemes.

We tested a control group of normally hearing subjects (NHS), and a group of deaf subjects before implantation at a time when they achieved a high level of performance in speech-reading. Deaf subjects were also tested at early and late periods after cochlear implantation when the role of audiovisual integration has been shown to increase (Rouger et al., 2007).

2. Materials and methods

2.1. Patients

To explore the role of gender differences during single-word speech-reading, we retrospectively analysed data from the study of Rouger et al. (2007). The deaf patients group comprised 60 females (aged 57 ± 17 (±S.D.) years) and 37 males (aged 55 ± 16 years). They received a CI after profound deafness (defined as a hearing loss of ≥90 dB) of diverse etiologies (meningitis, chronic otitis, otosclerosis and neurinoma) and durations (22 ± 15 years). Clinical implantation criteria included word and open-set sentence auditory-recognition scores below 30% under best-aided conditions with conventional acoustic hearing aids. All CI patients were recipients of a Nucleus (Cochlear) implant (CI-22 or CI-24) and used a range of different sound-coding strategies. All patients were similarly defined as profoundly deaf and the duration of deafness was not significantly different between men and women (18 and 22 years of deafness, respectively). Performance scores were collected during regular visits to the Ear, Nose and Throat department following a standard rehabilitation program. All participants gave their full-informed consent prior to their participation in this study in accordance with the Declaration of Helsinki (1975). For further details on patients see Rouger et al. (2007).

Data obtained from patients were analysed within three different testing periods defined according to the duration of experience with the cochlear implant. First, we pooled data from deaf patients tested once during a period of 18 months before cochlear implantation (41 women and 24 men, aged 55 ± 18 and 56 ± 16 years, respectively). A second set of data comprised the performance of CI patients tested once between 3 and 6 months after implantation (49 women and 31 men, aged 58 ± 16 and 54 ± 16 years, respectively). Lastly, we analysed the data obtained in CI patients tested once between 24 and 36 months after implantation (28 women and 18 men, aged 58 ± 16 and 51 ± 18 years, respectively). There was no significant gender difference in age within each group, neither between the three groups. Importantly, patients from the clinical database were implanted at different times and tested within different durations, so that approximately half of them were tested at all 3 time points. This explains the different number of subjects across the conditions.

In order to compare gender effects between word and phoneme speech-reading, we retrospectively analysed a second group of patients who participated in our previous study on the McGurk effect (Rouger, Fraysse, Deguine, & Barone, 2008). This group comprised 16 female (aged 48 ± 16 years) and 17 male CI subjects (aged 56 ± 14 years). All CI patients had a postlingually acquired deafness and were recipients of an unilateral Nucleus (Cochlear) implant (CI-22 or CI-24) using a range of different sound-coding strategies. Post-implantation time varied from 0 to 96 months (mean 21.4 ± 25.8 months). For further details on this group of patients, see Rouger et al. (2008).

2.2. Controls

For word recognition during speech-reading, we tested a group of normally hearing subjects, which comprised 20 females (aged 34 ± 10 years) and 22 males (aged 33 ± 11 years).

We analysed performance in phoneme comprehension during speech-reading in another group of normally hearing subjects, which comprised 20 females (aged 34 ± 14 years) and 19 males (aged 30 ± 12 years).

2.3. Stimuli and procedure

To assess word speech-reading performance, we used a recognition test including French disyllabic words derived from French word lists specifically designed for speech rehabilitation therapies. Stimuli were grouped into lists of 20 words each. Each list was equalized for syllabic structure (CV/CVC/CCV), language utilization frequency (Content et al., 1990), and anterior–posterior phonemic constitution; e.g., brevet (*certificate*), ruisseau (*stream*), bijou (*jewel*), danger (*danger*), sermon (*sermon*). The stimuli were recorded by a female French speech therapist, who pronounced each word with even intonation, tempo, and vocal intensity. The recordings were obtained in an anechoic chamber using a professional digital video camera with lights focused on the face so that minimal shadowing occurred.

For phoneme speech-reading, stimuli were meaningless VCV disyllables, where the vowel was always /a/ and the consonant was randomly chosen among the six

Author's personal copy

stops /p/ /b/ /t/ /d/ /k/ /g/, the two nasals /m/ /n/ and the six fricatives /f/ /v/ /s/ /z/ /h/ /j/. See Rouger et al. (2008) for the detailed description of these stimuli.

During the word speech-reading condition, a list of 20 words was presented in a random order in visual-only condition (image without sound). To test whether gender differences in visual recognition of words can be explained by the differences in auditory and audiovisual recognition, we also used auditory and audiovisual presentation of words. During the auditory condition, a second list of 20 words was presented in a random order in auditory-only condition (sound without image). During the audiovisual condition, a third list of 20 words was presented in a random order, with both image and sound. Words correctly repeated verbally by the subjects were treated as correct responses; the correct response score in % was calculated for each subject.

2.4. Analysis of the data

Several tests (Kolmogorov–Smirnov, Lilliefors, Bera-Jarque), rejected the null hypothesis of normality for the distribution of correct responses at the significance level of $p < 0.05$; deviation from the normal law was mostly in the asymmetry of empirical distributions. Taking into account the deviation from normality and the presence of repeated measures in our experimental protocol, we chose the bootstrap method for the statistical analysis (Carpenter & Bithell, 2000). The bootstrap simulation is also helpful to compensate for the unbalanced size of groups in our study. Using bootstrap, we obtained an equal number of 10,000 simulated observations per group in each condition.

The data for men and women was re-sampled 10,000 times. As a result, we obtained a distribution of 10,000 simulated observations for women and a distribution of 10,000 simulated observations for men in each condition, from which we obtained the means of the samples. We used bias-corrected and accelerated confidence intervals (Carpenter & Bithell, 2000) and estimated the significance of the mean correct response for men and women at the level of $p < 0.05$.

As this analysis demonstrated superior speech-reading performance in women, we calculated the relative gender difference in performance using the formula $(F_m - M_m)/F_m$ where F_m and M_m are the vectors obtained by re-sampling the original correct response scores 25 times for both females and males. Next, we re-sampled the relative difference 10,000 times to obtain the bias-corrected and accelerated confidence intervals ($p < 0.05$).

Calculations were performed in the R statistical environment (http://www.r-project.org) using the "boot" library.

3. Results

3.1. Speech-reading performance in normally hearing subjects and CI patients

For the speech-reading comprehension of words, normally hearing subjects displayed relatively low scores of 9.5 ± 1.7%. This contrasts with the higher speech-reading skills observed in the deaf patients at each testing period relative to cochlear implantation, which range from 29.7% (±3.5) before implantation, to 36.3% (±2.6) early after implantation and 37.7% (±3.5) at the longer interval post-implantation (Fig. 1). These non-overlapping confidence intervals demonstrate that normally hearing subjects are significantly less accurate for speech-reading than deaf patients tested before and after implantation. We observed a difference in performance levels between groups of patients tested before the implantation and between 3 and 6 months after implantation, but we did not find any statistical difference when comparing the two post-implantation groups (3–6 vs. 24–36 months post-implantation). In our previous study (Rouger et al., 2007), we reported a tendency for an increase in speech-reading performance from the period of pre-implantation to the first year of experience with the neuroprosthesis, although a paired comparison analysis confined only to patients who underwent the entire follow-up did not reveal a significant difference. Nevertheless, these data suggest that the high speech-reading proficiency developed by deaf patients may slightly improve following the reception of a cochlear implant.

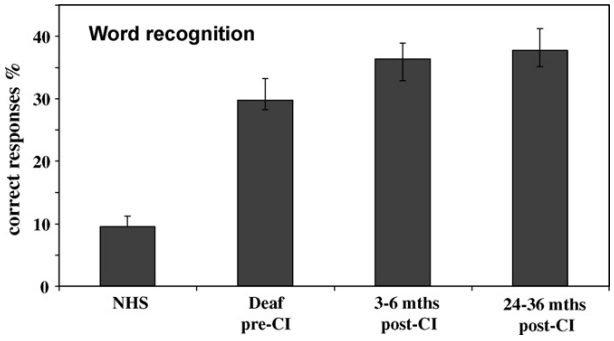


Fig. 1. Global effect for speech-reading word recognition in normally hearing subjects (NHS) and deaf patients before (deaf pre-CI), early after (3–6 months), or late after (24–36 months) cochlear implantation (CI). Error bars represent bootstrap bias-corrected and accelerated confidence intervals ($p < 0.05$). Note the superior performance of deaf patients in speech-reading skills compared to NHS.

For phonemes, correct responses were 26.6 ± 6.7% in normally hearing subjects and 27.7 ± 4.2% in patients with cochlear implants (confidence intervals at the level of $p < 0.05$). As confidence intervals overlap, there is no significant difference between normally hearing subjects and patients.

3.2. Gender effect for word recognition

Further analyses were carried out to examine the performance levels of subjects according to the gender. Strong differences were observed in both the control and deaf patient groups (Fig. 2).

In normally hearing subjects, the performance for words was 12.3 ± 2.8% for females and 6.8 ± 1.3% for males. This difference was statistically significant, leading to a relative superiority of 44.5 ± 0.5% in women compared to men.

In deaf patients, the gender difference was also present when patients were tested before the implantation. The percentage of correct responses was 33.3 ± 4.1% for females and only 23.5 ± 5.2% for males. In this case, the relative superiority of women in speech-reading is 29.2 ± 2.5% compared to men.

When patients were tested during the first 3–6 months post-implantation, we observed a tendency for an improvement in speech-reading performance in both genders compared to that observed before implantation, although this tendency remains above the significance level. Correct responses for words were 39.5 ± 3.2% for females and 31.3 ± 3.7% for males. Consequently, women are significantly better speech-readers than men within the first month post-implantation with a relative superiority of 21 ± 2.6%.

Later, after 2 years of auditory recuperation with the implant, we observed a specific increase in performance for the male group compared to the pre-implantation period (37.5% vs. 23.5%, respectively) while the performance in women did not show any statistically significant improvement. This improvement led to similar performance in both genders late after the implantation, with correct word recognition of 37.9 ± 5.5% for females and 37.5 ± 4.7% for males.

Thus, to summarize, we observed a significant gender difference in speech-reading abilities, with higher performance levels in women (Fig. 3). This difference is most pronounced in normally hearing subjects. In cochlear implanted deaf patients, this gender difference is dependent on the duration of experience with the neuroprosthesis, being large in deaf non-implanted subjects, reduced in the first 6 months post-implantation and not significantly different from zero in patients implanted for a duration since more than 2 years.

It is now well established that the loss of one sensory modality induces a cross-modal compensation, leading to improved abilities in another sensory channel (Bavelier & Neville, 2002; Bavelier et al., 2006). All patients were classified as "profoundly deaf" as they were presenting a hearing loss greater than 90 dB. However the

Author's personal copy

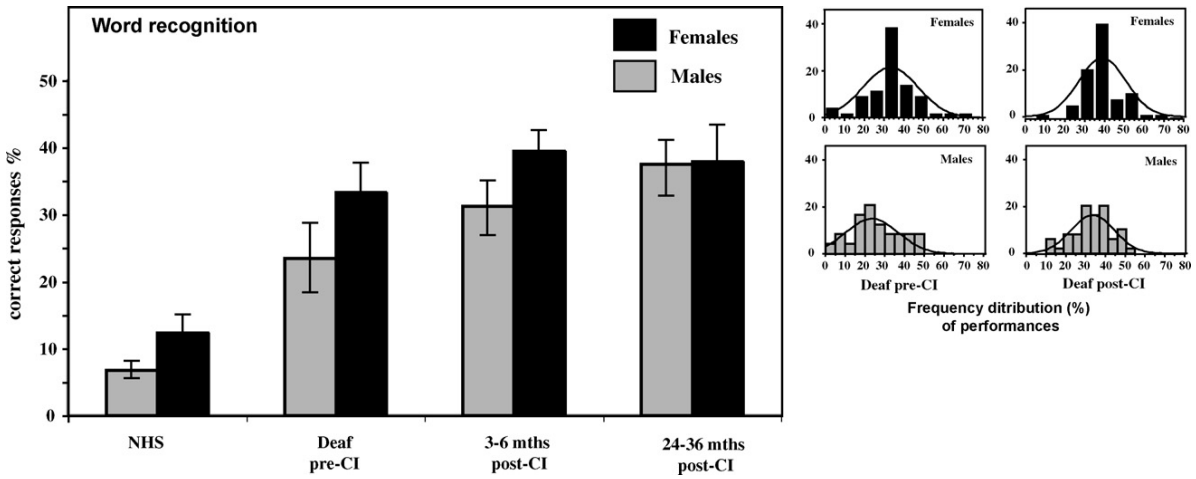


Fig. 2. Gender effect for speech-reading word recognition. With the exception of the group of experienced CI patients (24–36 months post-CI), we observed superior performance in women compared to men. The right graph shows the frequency distribution (in %) of performance of male and female CI users separately. Axes are scaled to the same ranges of values. The approximate normal distribution curve is presented for comparison. Same conventions as in Fig. 1.

auditory performance in word recognition can greatly vary across deaf patients (range before implantation: 0–30% with conventional external hearing aids). Therefore, the superior performance in speech-reading observed in female deaf patients could possibly be due to a higher degree of hearing loss. While this hypothesis does not explain the gender differences revealed in normally hearing subjects, we examined whether this could be the cause in the deaf patients by analysing their performance in auditory word recognition. We did not find any significant differences for auditory performance when comparing men and women at any testing period, before or after cochlear implantation. Before implantation, the performance is very low in both groups of patients (women: $10.6 \pm 4.1\%$; men: $10.3 \pm 5.9\%$) and then increases progressively during the first months after implantation (women: $77.8 \pm 7\%$; men: $66.5 \pm 11\%$) to reach a plateau after 1–2 years (women: $86.0 \pm 10.2\%$; men: $82.5 \pm 14.6\%$) as previously reported (Rouger et al., 2007).

Finally, we examined gender differences in the bimodal audiovisual condition. Here, word recognition performance was much higher in all groups of patients, ranging from $57.9 \pm 5.8\%$ (women) and $51.3 \pm 9.3\%$ (men) before implantation to nearly perfect scores after implantation: $97.0 \pm 1.7\%$ (women) and $94.8 \pm 3.2\%$ (men) early after implantation; $97.8 \pm 1.9\%$ (women) and $99.2 \pm 0.5\%$ (men) late after implantation. Statistical analysis did not reveal any gender effect for audiovisual word recognition in any of these three groups of patients. In addition and to conclude, as the etiology and duration of deafness did not differ between women and men (see Section 2), the gender differences we observed are very unlikely to result from differences in the severity or the type of deafness.

3.3. Gender effect for phoneme recognition

Lastly, we analysed the speech-reading performance in normally hearing and cochlear implanted subjects using phonemes. For patients, phoneme recognition was more difficult than word recognition (see Rouger et al., 2008), a finding that can be explained by the absence of lexical cues. In the normally hearing group, correct responses for phonemes were $25.6 \pm 7.9\%$ for females and $27.5 \pm 12.5\%$ for males ($p < 0.05$ for the confidence intervals), with no significant gender difference.

In cochlear implanted patients, independent of the period following the implantation, correct responses for phonemes were similar to those obtained in controls, at a level of $28.9 \pm 5.5\%$ for the females and $26.5 \pm 7\%$ for the males. Again, as in the normally hearing group, we did not observe any significant gender difference (Fig. 4).

4. Discussion

Our results demonstrate that women are better speech-readers than men for single-word recognition except for experienced

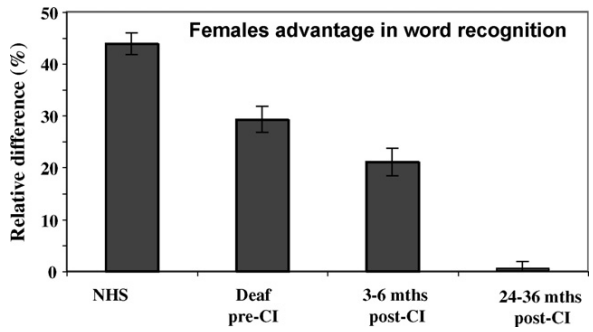


Fig. 3. Relative superiority of women for speech-reading skills compared to men in all groups of subjects. Female advantage for speech-reading is the highest in normally hearing subjects. Same conventions as in Fig. 1.

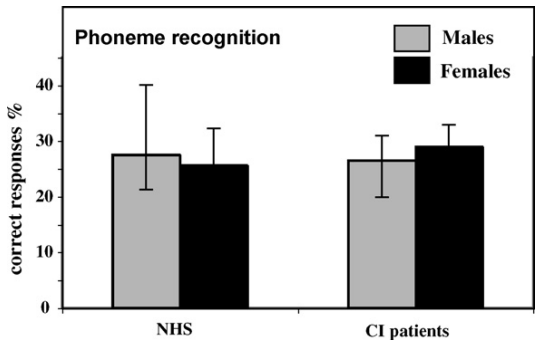


Fig. 4. Speech-reading performance for phoneme recognition. No gender difference can be observed in normally hearing subjects neither in CI deaf patients. Same conventions as in Fig. 1.

Author's personal copy

cochlear implant users. This difference is observed both in normally hearing subjects who are poor speech-readers, as well as in deaf people who have developed a strong proficiency in speech-reading to compensate for their deafness. The relative superiority of women is different across the groups and conditions, being the highest in normally hearing subjects, smaller in deaf subjects, and absent in implanted deaf subjects, who have recovered auditory speech comprehension.

4.1. Gender differences for visual speech

A small number of previous studies have reported that women are more strongly affected by visual speech information during speech comprehension. For example, using the ambiguous McGurk effect paradigm, the audiovisual percept has been shown to be more strongly influenced by the visual cues in women compared to men, especially during brief stimulus presentation (Aloufy et al., 1996; Irwin et al., 2006). However, for visual speech comprehension *per se*, reports on gender differences are contradictory, most probably because of the variety of protocols used, which have ranged from syllables to words and sentence comprehension tests (Daly, Bench, & Chappel, 1997; Johnson & Leslie, 1988; Watson et al., 1996). The results of the present study show a consistent superiority of women for word but not for phoneme recognition, and thus confirm the hypothesis that women are better speech-readers for continuous speech fragments. This calls into question the proposition (Jaeger et al., 1998) that the more task demands increase, the more gender differences can be expected. Indeed, phoneme perception seems to be more difficult because there are less contextual clues than in the case of words. The perception of sounds in a known word is facilitated by phonological, morphological and semantic predictive coding based on the word representation in linguistic memory, which can be activated by the lacking linguistic information as suggested by the activation-verification model of continuous speech processing (Strelnikov, 2008). We suggest that the key aspect for gender differences could be the necessity of predictive integrations.

In their review of neuroimaging data, Kansaku and Kitazawa (2001) use a more general assertion, claiming that gender differences in auditory and written speech perception are more likely to occur when subjects process global speech structures but not single words. The authors localized these differences within the superior and middle temporal regions, with men having more left-lateralized activations. This provides another line of support for the suggestion that gender differences are related to the continuity of speech, i.e. to the higher involvement of predictive and integrative strategies during speech processing. Contrary to auditory or written speech perception, speech-reading is much poorer in linguistic clues, thus its processing involves predictive and integrative strategies to a greater extent than usual speech. This explains why we observed gender differences in the present study, while they are absent when using auditory and written speech (Kansaku & Kitazawa, 2001). During speech-reading, the more subjects have to guess the speech stimuli while trying to integrate the insufficient perceptual clues, the more the superiority of women can be expected. The superiority of woman may be based on the higher involvement of the right hemispheric structures. To our knowledge, there is no neuroimaging study of gender differences during speech-reading of words, however during speech-reading of a predefined list of numbers, women demonstrated more activation than men in the right homologue of Broca area (Ruytjens, Albers, van Dijk, Wit, & Willemsen, 2006). Taking into account our suggestion about peculiarities of speech processing strategies in women, we can predict that at the sentence level the superiority of women will be preserved or even greater. It will be important to examine this in future studies on speech-reading at the behavioural level, as well as with neuroimaging and electrophysiological techniques.

Several mechanisms are implicated in speech-reading skills from sensory, memory, cognitive and linguistic processes (see Gailey, 1987 for review). In the first stage, visual analysis of the face (lips, tongue, teeth and muscle movements) must be performed in order to extract phonemic information. Thus, the female advantage we observed for speech-reading could find its origin in differences at the sensory level with the involvement of attentional mechanisms. In the visual modality, behavioural studies have shown some gender differences in spatial tasks (see Voyer, Voyer, & Bryden, 1995 for the review) including extraction of cues from visual scenes (Barkley & Gabriel, 2007). These differences could account for some gender dissimilarities during processing of the visual cues provided by lip movements. On another hand, women are more attentive gazers (Deaner, Shepherd, & Platt, 2007) and are better than men at recognising facial expressions (Rahman, Wilson, & Abrahams, 2004; Thayer & Johnsen, 2000), so that visually based communicative support is more developed in women. However, it should be noted that face processing and speech-reading mechanisms are functionally dissociated in the human brain (Campbell, Landis, & Regard, 1986). Furthermore, the hypothesis of higher visual skills in women does not concur with the absence of any observed gender difference during speech-reading of isolated phonemes.

It is important to note that gender differences in cognitive functions are not only related to speech and facial processing but rather constitute a general biological phenomenon: behavioural studies have reported gender differences across some visual tasks, although only few have described such gender differences when the auditory function is involved (van Kesteren & Wiersinga-Post, 2007; but see Fink, Churan, & Wittmann, 2005). It is noteworthy that gender differences are frequently observed in different types of visuospatial tasks, such as mental and spatial rotation (Parsons et al., 2004), cognitive mapping (MacFadden, Elias, & Saucier, 2003; Rahman, Abrahams, & Jussab, 2005), orientation and navigation strategies (Rahman, Andersson, & Govier, 2005), visuospatial attention (Robinson & Kertzman, 1990), and vocal emotion processing (Schirmer, Striano, & Friederici, 2005). The social and biological bases of these cognitive differences remain to be clarified. One may wonder whether speech-reading, at least partly, also falls into this category of cognitive tasks and whether there are common brain mechanisms underlying those gender differences.

4.2. Neural bases of the gender difference in speech-reading

As previously described, visuospatial skills, like mental rotation, are one of the few cognitive domains for which gender differences have been formally established (see Voyer et al., 1995 for review). In conjunction with these behavioural differences, brain imaging studies have reported cerebral activation patterns that differ between male and female subjects (Gron, Wunderlich, Spitzer, Tomczak, & Riepe, 2000). More recently, gender differences have been observed when subjects are performing motor (Lissek et al., 2007) and auditory tasks (Ruytjens et al., 2007). How does this relate to our results on gender differences in speech-reading?

The cortical regions involved in speech-reading have recently been explored in several studies using brain imaging procedures (Bernstein et al., 2002; Paulesu et al., 2003; Skipper, van Wassenhove, Nusbaum, & Small, 2007) and through analysing the effects of restricted brain lesions (Campbell, Zihl, Massaro, Munhall, & Cohen, 1997; Hamilton, Shenton, & Coslett, 2006). Speech-reading is accomplished through a complex anatomical network, including regions from visual movement areas to the auditory cortex and to the inferior frontal and premotor areas (see Campbell, 2008). While it is now well accepted that the auditory associative

Author's personal copy

areas are activated during speech-reading (Bernstein et al., 2002; Paulesu et al., 2003), the involvement of the primary auditory cortex (Calvert et al., 1997; MacSweeney et al., 2000; Pekkola et al., 2005) is still subject to controversy in normal hearing subjects. However, this may be due to mostly technical problems, and spatial resolution may be not sufficient to distinguish exactly between primary and associative auditory areas.

Our present data could help to clarify how the neural networks involved in speech-reading may differ between males and females. This issue has recently been addressed in a PET study which showed that during a speech-reading task, female subjects present a broader spread of cortical activation than male subjects (Ruytjens et al., 2006). In particular, activation of the speech-reading areas was much more left-lateralized in the male than in the female group. It is relevant to note that this gender difference was observed during a "simplified" speech-reading task (a closed list of numbers) for which the two groups obtained similar levels of performance. We would predict that for a more complex speech-reading task using randomly chosen words, this gender difference in brain activations would be preserved and possibly even accentuated.

As mentioned in Section 1, when speech processing involves predictive and integrative cognitive processes, more bilateral and right hemispheric activations can be expected. Our results suggest that with training following cochlear implantation, men can learn to process speech using an integrative and predictive strategy, resulting in a reduction of the speech-reading advantage found for women. The effectiveness of learning in men may be explained by the following consideration: since men use more left-lateralized and localized neural systems for speech processing, there is greater potential for the recruitment of the more widespread neural networks when adapting to the new auditory input with the implant. Under normal conditions, women already use more symmetrical and diffuse neural resources (not only for speech, but for face processing as well) (Proverbio, Brignone, Matarazzo, Del Zotto, & Zani, 2006). For speech-reading, normally hearing women use a more extended network than men (Ruytjens et al., 2006). Thus, the potential for the recruitment of more widespread neural resources in women may be smaller.

4.3. Speech-reading in deaf and cochlear implanted patients: implication for cortical cross-modal reorganization

Cross-modal compensation has been shown to follow long periods of visual or auditory deprivation (Bavelier & Neville, 2002; Bavelier et al., 2006). In the case of profound deafness, the acquisition of speech-reading skills is one of the sensory substitution strategies developed by deaf patients in order to recover a degree of speech comprehension (Grant, Walden, & Seitz, 1998; Kaiser, Kirk, Lachs, & Pisoni, 2003; Summerfield, 1992; Tyler, Parkinson, Woodworth, Lowder, & Gantz, 1997). In our recent studies, we have demonstrated that there is an increased capacity for speech-reading a few months after the onset of deafness and that deaf patients maintain a high level of speech-reading for years following cochlear implantation (Rouger et al., 2007). In addition, in case of an ambiguity in audiovisual speech, cochlear implanted patients are more strongly influenced by the visual information (Rouger et al., 2008). We have hypothesized that CI users maintain a high level of speech-reading performance because they are very sensitive to noisy environment (Fu, Shannon, & Wang, 1998), so that this strategy is crucial to compensate for deficiencies of speech comprehension in the auditory modality. However, several studies have reported contradictory results. While some studies have observed a progressive but strong increase in speech-reading skills following implantation (Giraud, Price, Graham, & Frackowiak, 2001; but see Giraud, Truy, & Frackowiak, 2001), others have reported only mild (Bergeson, Pisoni, & Davis, 2005) or no changes (Rouger et al., 2007) in visual performance. These discrepancies could result from the material used for testing the patients (words, phonemes recognition in sentences, etc.), but could also be due to the proportion of male and female subjects as, according to the present results, changes in speech-reading performance after implantation depend on the individual's gender.

A resulting question concerns the mechanisms through which visual speech recognition in deaf males increases after implantation and why this improvement becomes more dependent on the recovery of auditory functions than in women. We propose that this improvement results from an involvement of multisensory integration in unisensory perceptual learning. In support of this, it has been shown that cross-modal reorganization in CI users is strongly related to audio–visual coupling (Doucet, Bergeron, Lassonde, Ferron, & Lepore, 2006). Furthermore, recent psychophysical studies have shown that audio–visual training can increase the rate of learning and can improve perceptual performance in the auditory modality alone (Lippert, Logothetis, & Kayser, 2007; Seitz, Kim, & Shams, 2006). Multisensory training has also been successfully employed as a clinical rehabilitation strategy in patients suffering from sensory neglect (Frassinetti, Pavani, & Ladavas, 2002). It is likely that Hebbian mechanisms contribute to the success of these multisensory training strategies (Rauschecker, 1991) and they underlie the improved performance of male patients for speech-reading. The visual information provided by speech-reading is ambiguous, and some phonemes cannot be distinguished from the visual cues alone (Walden, Prosek, Montgomery, Scherr, & Jones, 1977). We hypothesize that in deaf male patients, the progressive recovery of auditory and audiovisual speech comprehension following cochlear implantation provides a strong positive feedback that consolidates the "decoding" of visual cues into speech information and results in an increase in speech-reading performance. This is supported by the observation that the audiovisual performance in men constantly increases with time, while for women it remains the same in the early and late periods following implantation. It would appear that ceiling performance for speech-reading in women is reached prior to cochlear implantation, and so further improvements cannot be observed following implantation.

In deaf patients, cross-modal compensations phenomena have been linked to cortical functional reorganization (Bavelier & Neville, 2002). For example, activations within the auditory areas have been found when deaf subjects are viewing visual stimuli that contain linguistic information, such as speech-reading stimuli and sign language (Nishimura et al., 1999; Petitto et al., 2000). This cross-modal reorganization might affect differently the auditory areas depending on deafness duration (Lee et al., 2007) and severity of hearing loss (Lambertz, Gizewski, de Greiff, & Forsting, 2005). The present data suggest that functional reorganization that can be observed in cortical areas related to speech processing (Capek et al., 2008; Giraud, Price, Graham, & Frackowiak, 2001; Giraud, Price, Graham, Truy et al., 2001; Lee et al., 2007) probably follows a different time-course in men and women. After several years of experiences with the implant, male and female CI patients reach similar levels of recovery of speech perception; yet our data suggest that the synergy between auditory and visual speech processing is involved within different time scales in this progressive recovery.

4.4. Conclusions

Women are better speech-readers for words – but not for phonemes – in normally hearing controls and deaf patients before and early after cochlear implantation, but achieve similar performance to men during the late period after cochlear implantation. This is in line with our hypothesis formulated in Section 1 that

Author's personal copy

women use predictive and integrative strategies to a greater extent during speech processing, especially when it is needed for better adaptation. However, when necessary for adaptation, the ability of such strategies may also improve in men nearly reaching with time the level of women. Several studies have attempted to tackle the issue concerning the origins of gender differences for speech-reading, specifically whether they reflect nature or nurture effects (Kass & Ahlers, 1998; Silverman, Phillips, & Silverman, 1996). The large difference in speech-reading skills between males and females is probably more indicative of a gender difference in communication and interaction strategies rather than of some biological constraints. Such interpretation is supported by the absence of such gender difference in experienced cochlear implanted patients. Furthermore, our results suggest that the cortical mechanisms of cross-modal plasticity following hearing loss (Giraud, Price, Graham, & Frackowiak, 2001) may have different time-course in males and females. Functional remodelling of the brain caused by the long-term sensory deprivation might thus have different effects on the cortical networks in each gender.

Acknowledgments

We thank the cochlear implanted and normally hearing subjects for their participation in this study, Marie-Laurence Laborde for help in collecting the data, and Jessica Foxton for critical correction of the manuscript. This work was supported by grants from the CNRS Atip+ program (KS), the Fondation pour la Recherche Médicale (SL and JR) and the recurrent funding of the CNRS (PB and OD).

References

Aloufy, S., Lapidot, M., & Myslobodskym, M. (1996). Differences in susceptibility to the "blending illusion" among Native Hebrew and English speakers. *Brain and Language*, *53*(1), 51–57.

Barkley, C. L., & Gabriel, K. I. (2007). Sex differences in cue perception in a visual scene: Investigation of cue type. *Behavioral Neuroscience*, *121*(2), 291–300.

Bavelier, D., & Neville, H. J. (2002). Cross-modal plasticity: Where and how? *Nature Reviews*, *3*(6), 443–452.

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., et al. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *Journal of Neuroscience*, *20*(17), RC93.

Bavelier, D., Dye, M. W., & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, *10*(11), 512–518.

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing*, *26*(2), 149–164.

Bernstein, L. E., Auer, E. T., Jr., & Tucker, P. E. (2001). Enhanced speechreading in deaf adults: Can short-term training/practice close the gap for hearing adults? *Journal of Speech, Language and Hearing Research*, *44*(1), 5–18.

Bernstein, L. E., Auer, E. T., Jr., Moore, J. K., Ponton, C. W., Don, M., & Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *Neuroreport*, *13*(3), 311–315.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science (New York, NY)*, *276*(5312), 593–596.

Campbell, R. (2008). The processing of audio–visual speech: Empirical and neural bases. *Philosophical Transactions of the Royal Society of London*, *363*(1493), 1001–1010.

Campbell, R., Landis, T., & Regard, M. (1986). Face recognition and lipreading. A neurological dissociation. *Brain*, *109*(Pt 3), 509–521.

Campbell, R., Zihl, J., Massaro, D., Munhall, K., & Cohen, M. M. (1997). Speechreading in the akinetopsic patient, L.M. *Brain*, *120*(Pt 10), 1793–1803.

Capek, C. M., Macsweeney, M., Woll, B., Waters, D., McGuire, P. K., David, A. S., et al. (2008). Cortical circuits for silent speechreading in deaf and hearing people. *Neuropsychologia*, *46*(5), 1233–1241.

Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, *19*(9), 1141–1164.

Collignon, O., Renier, L., Bruyer, R., Tranduy, D., & Veraart, C. (2006). Improved selective and divided spatial attention in early blind subjects. *Brain Research*, *1075*(1), 175–182.

Content, A., Mousty, P., & Radeau, M. (1990). Brulex. une base de données lexicales informatisée pour le français écrit et parlé. *Annals of Psychology*, *90*, 551–566.

Daly, N., Bench, J., & Chappel, H. (1997). Gender differences in visual intelligibility. *Journal of the Academy of Rehabilitative Audiology*, *30*, 63–76.

Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters*, *3*(1), 64–67.

Doucet, M. E., Bergeron, F., Lassonde, M., Ferron, P., & Lepore, F. (2006). Cross-modal reorganization and speech perception in cochlear implant users. *Brain*, *129*(Pt 12), 3376–3383.

Fine, I., Finney, E. M., Boynton, G. M., & Dobkins, K. R. (2005). Comparing the effects of auditory deprivation and sign language within the auditory and visual cortex. *Journal of Cognitive Neuroscience*, *17*(10), 1621–1637.

Fink, M., Churan, J., & Wittmann, M. (2005). Assessment of auditory temporal-order thresholds—A comparison of different measurement procedures and the influences of age and gender. *Restorative Neurology and Neuroscience*, *23*(5–6), 281–296.

Finney, E. M., Fine, I., & Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience*, *4*(12), 1171–1173.

Frassinetti, F., Pavani, F., & Ladavas, E. (2002). Acoustical vision of neglected stimuli: Interaction among spatially converging audiovisual inputs in neglect patients. *Journal of Cognitive Neuroscience*, *14*(1), 62–69.

Fu, Q. J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *104*(6), 3586–3596.

Gailey, L. (1987). Psychological parameters of lip-reading skill. In R. Dodd & B. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 115–141). Erlbaum: Hillsdale, NJ.

Giraud, A. L., Price, C. J., Graham, J. M., & Frackowiak, R. S. (2001). Functional plasticity of language-related brain areas after cochlear implantation. *Brain*, *124*(Pt 7), 1307–1316.

Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., & Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*(3), 657–663.

Giraud, A. L., Truy, E., & Frackowiak, R. (2001). Imaging plasticity in cochlear implant patients. *Audiology and Neurootology*, *6*(6), 381–393.

Goldreich, D., & Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *Journal of Neuroscience*, *23*(8), 3439–3445.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Neuropsychology: Pitch discrimination in the early blind. *Nature*, *430*(6997), 309.

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory–visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory–visual integration. *The Journal of the Acoustical Society of America*, *103*(5), 2677–2690.

Green, K. M., Julyan, P. J., Hastings, D. L., & Ramsden, R. T. (2005). Auditory cortical activation and speech perception in cochlear implant users: Effects of implant experience and duration of deafness. *Hearing Research*, *205*(1–2), 184–192.

Gron, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: Gender-different neural networks as substrate of performance. *Nature Neuroscience*, *3*(4), 404–408.

Hamilton, R. H., Shenton, J. T., & Coslett, H. B. (2006). An acquired deficit of audiovisual speech processing. *Brain and Language*, *98*(1), 66–73.

Irwin, J. R., Whalen, D. H., & Fowler, C. A. (2006). A sex difference in visual influence on heard speech. *Perception & Psychophysics*, *68*(4), 582–592.

Ito, K., Momose, T., Oku, S., Ishimoto, S., Yamasoba, T., Sugasawa, M., et al. (2004). Cortical activation shortly after cochlear implantation. *Audiology and Neurootology*, *9*(5), 282–293.

Jaeger, J. J., Lockwood, A. H., Van Valin, R. D., Jr., Kemmerer, D. L., Murphy, B. W., & Wack, D. S. (1998). Sex differences in brain regions activated by grammatical and reading tasks. *Neuroreport*, *9*(12), 2803–2807.

Johnson, F. M., & Leslie, H. H. (1988). Sex differences in lipreading. *Bulletin of Psychonomic Society*, *26*(2), 106–108.

Kaiser, A. R., Kirk, K. I., Lachs, L., & Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language and Hearing Research*, *46*(2), 390–404.

Kansaku, K., & Kitazawa, S. (2001). Imaging studies on sex differences in the lateralization of language. *Neuroscience Research*, *41*(4), 333–337.

Kansaku, K., Yamaura, A., & Kitazawa, S. (2000). Sex differences in lateralization revealed in the posterior language areas. *Cerebral Cortex*, *10*(9), 866–872.

Kass, S. J., & Ahlers, R. H. (1998). Eliminating gender differences through practice in applied visual spatial task. *Human Performance*, *11*(4), 337–349.

Kitazawa, S., & Kansaku, K. (2005). Sex difference in language lateralization may be task-dependent. *Brain*, *128*(Pt 5) (E30; author reply E31)

Lambertz, N., Gizewski, E. R., de Greiff, A., & Forsting, M. (2005). Cross-modal plasticity in deaf subjects dependent on the extent of hearing loss. *Brain Research, Cognitive Brain Research*, *25*(3), 884–890.

Lee, H. J., Truy, E., Mamou, G., Sappey-Marinier, D., & Giraud, A. L. (2007). Visual speech circuits in profound acquired deafness: A possible role for latent multimodal connectivity. *Brain*, *130*(Pt 11), 2929–2941.

Lippert, M., Logothetis, N. K., & Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain Research*, *1173*, 102–109.

Lissek, S., Hausmann, M., Knossalla, F., Peters, S., Nicolas, V., Gunturkun, O., et al. (2007). Sex differences in cortical and subcortical recruitment during simple and complex motor control: An fMRI study. *NeuroImage*, *37*(3), 912–926.

MacFadden, A., Elias, L., & Saucier, D. (2003). Males and females scan maps similarly, but give directions differently. *Brain and Cognition*, *53*(2), 297–300.

Author's personal copy

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., et al. (2000). Silent speechreading in the absence of scanner noise: An event-related fMRI study. *Neuroreport*, *11*(8), 1729–1733.
Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., et al. (1999). Sign language 'heard' in the auditory cortex. *Nature*, *397*(6715), 116.
Ortigue, S., Thut, G., Landis, T., & Michel, C. M. (2005). Time-resolved sex differences in language lateralization. *Brain*, *128*(Pt 5) (E28; author reply E29).
Parsons, T. D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J. G., et al. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, *42*(4), 555–562.
Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., et al. (2003). A functional-anatomical model for lipreading. *Journal of Neurophysiology*, *90*(3), 2005–2013.
Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I. P., Mottonen, R., Tarkiainen, A., et al. (2005). Primary auditory cortex activation by visual speech: An fMRI study at 3 T. *Neuroreport*, *16*(2), 125–128.
Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., & Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13961–13966.
Phillips, M. D., Lowe, M. J., Lurito, J. T., Dzemidzic, M., & Mathews, V. P. (2001). Temporal lobe activation demonstrates sex-based differences during passive listening. *Radiology*, *220*(1), 202–207.
Proverbio, A. M., Brignone, V., Matarazzo, S., Del Zotto, M., & Zani, A. (2006). Gender differences in hemispheric asymmetry for face processing. *BMC Neuroscience*, *7*, 44.
Putzar, L., Goerendt, I., Lange, K., Rosler, F., & Roder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience*, *10*(10), 1243–1245.
Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition*, *54*(3), 179–185.
Rahman, Q., Abrahams, S., & Jussab, F. (2005). Sex differences in a human analogue of the Radial Arm Maze: The "17-Box Maze Test". *Brain and Cognition*, *58*(3), 312–317.
Rahman, Q., Andersson, D., & Govier, E. (2005). A specific sexual orientation-related difference in navigation strategy. *Behavioral Neuroscience*, *119*(1), 311–316.
Rauschecker, J. P. (1991). Mechanisms of visual plasticity: Hebb synapses, NMDA receptors, and beyond. *Physiological Review*, *71*(2), 587–615.
Robinson, D. L., & Kertzman, C. (1990). Visuospatial attention: Effects of age, gender, and spatial reference. *Neuropsychologia*, *28*(3), 291–301.
Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, *400*(6740), 162–166.
Rossell, S. L., Bullmore, E. T., Williams, S. C., & David, A. S. (2002). Sex differences in functional brain activation during a lexical visual field task. *Brain and Language*, *80*(1), 97–105.
Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(17), 7295–7300.
Rouger, J., Fraysse, B., Deguine, O., & Barone, P. (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Research*, *1188*, 87–99.
Ruytjens, L., Albers, F., van Dijk, P., Wit, H., & Willemsen, A. (2006). Neural responses to silent lipreading in normal hearing male and female subjects. *The European Journal of Neuroscience*, *24*(6), 1835–1844.
Ruytjens, L., Georgiadis, J. R., Holstege, G., Wit, H. P., Albers, F. W., & Willemsen, A. T. (2007). Functional sex differences in human primary auditory cortex. *European Journal of Nuclear Medicine and Molecular Imaging*, *34*(12), 2073–2081.
Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, *16*(6), 635–639.
Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology*, *16*(14), 1422–1427.
Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature*, *373*(6515), 607–609.
Silverman, I., Phillips, K., & Silverman, L. K. (1996). Homogeneity of effect sizes for sex across spatial tests and cultures: Implications for hormonal theories. *Brain and Cognition*, *31*(1), 90–94.
Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, *17*(10), 2387–2399.
Sommer, I. E., Aleman, A., Bouma, A., & Kahn, R. S. (2004). Do women really have more bilateral language representation than men? A meta-analysis of functional imaging studies. *Brain*, *127*(Pt 8), 1845–1852.
Strelnikov, K. (2008). Activation-verification in continuous speech processing. Interaction of cognitive strategies as a possible theoretical approach. *Journal of Neurolinguistics*, *21*, 1–17.
Summerfield, Q. (1992). Lipreading and audio–visual speech perception. *Philosophical Transactions of the Royal Society of London*, *335*(1273), 71–78.
Thayer, J. F., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology*, *41*(3), 243–246.
Tyler, R. S., Parkinson, A. J., Woodworth, G. G., Lowder, M. W., & Gantz, B. J. (1997). Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *The Journal of the Acoustical Society of America*, *102*(1), 508–522.
van Kesteren, M. T., & Wiersinga-Post, J. E. (2007). Auditory temporal-order thresholds show no gender differences. *Restorative Neurology and Neuroscience*, *25*(2), 119–122.
Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250–270.
Walden, B. E., Prosek, R. A., Montgomery, A. A., Scherr, C. K., & Jones, C. J. (1977). Effects of training on the visual recognition of consonants. *Journal of Speech, Language and Hearing Research*, *20*(1), 130–145.
Watson, C. S., Qiu, W. W., Chamberlain, M. M., & Li, X. (1996). Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition. *The Journal of the Acoustical Society of America*, *100*(2 Pt 1), 1153–1162.

Index

A

affriquées, 50
aires de Brodmann, 341
alvéolaires, 47
API (alphabet phonétique international), 48
appareil phonatoire, 42
arrondissement, 49
articulateur, 47
atlas de Talairach et Tournoux, 348
audition, 17
    neurologie, 29
    perception, 26

C

canal cochléaire, 20
cavité buccale, 46
cavité labiale, 46
cavité pharyngale, 46
cavités nasales, 46
cellules ciliées, 22
cerveau, 339
cervelet, 340
cochlée, 20
colliculus inférieur, 30
complexe olivaire supérieur, 29
consonnes, 47, 50
    caractéristiques acoustiques, 54
    caractéristiques articulatoires, 50
constrictives, 50
cordes vocales, 44
corps genouillé médian, 30
cortex auditif associatif, 31
cortex auditif primaire, 31
cortex auditif secondaire, 31
cortex cérébral, 341

D

dentales, 47
décibels (dB), 27

E

effet McGurk, 65
enclume, 18
endolymphe, 20
espace supralaryngé, 45
étrier, 18

F

fenêtre ovale, 18
fenêtre ronde, 18
filtres auditifs, 27
formants, 53
fosses nasales, 46
fricatives, 50
fréquence fondamentale, 52

G

ganglion spiral, 29
glides, 47, 51
gyrus, 341
gyrus de Heschl, 31
gyrus temporal transverse, 31

I

implant cochléaire, 145
intégration multisensorielle, 109

L

labiales, 47
labialisation, 49
langage des signes, 140
langue, 46
larynx, 43
latérales, 51
lèvres, 46
lieu d'articulation, 47
liquides, 56
LPC (langage parlé complété), 141

M

marteau, 18
membrane basilaire, 20
membrane de Reissner, 20
membrane tectoriale, 22
MNI (Montreal Neurological Institute), 348
mode articulatoire, 46
modiolus, 20
mémoire auditive, 34

N

nasales, 50, 51
nasalisation, 49, 51
neuroanatomie, 339

non voisés, 47
notation phonétique, 48
noyaux cochléaires, 29
noyaux du lemnisque latéral, 29

O

occlusives, 50
orales, 50, 51
oreille externe, 18
oreille interne, 20
oreille moyenne, 18
organe de Corti, 22
osselets, 18
otosclérose, 134
otospongiose, 134
ototoxicité, 137

P

palatales, 47
perception de la parole
    chez l'implanté cochléaire, 192
    chez le sujet normal, 59
    chez le sujet sourd, 137
périlymphe, 20
pharynx, 46
phonation, 46
phonème, 46
phonétique, 48
planum temporale, 31
plosives, 50
production de la parole, 41

R

rampe tympanique, 20
rampe vestibulaire, 20

S

semi-consonnes, 47
semi-occlusives, 50
semi-voyelles, 47
sonantes, 50
spectrogramme, 53
SPM (statistical parametric mapping), 228, 348
sulcus, 341
surdité, 127
surdité brusque, 128
surdité centrale, 128
surdité congénitale, 128
surdité de perception, 128
surdité de transmission, 128
surdité postlinguale, 128
surdité progressive, 128
surdité prélinguale, 128
système auditif central, 29
système auditif périphérique, 18

T

test de Rinne, 130
test de Weber, 130
thalamus, 340
tonotopie, 32
tronc cérébral, 340
tube cochléaire, 20
tunnel de Corti, 22
tympan, 18

U

uvulaires, 47

V

vélaires, 47
velum, 46
vibrantes, 51
voile du palais, 46
voisés, 47
voyelles, 47, 48
    caractéristiques acoustiques, 54
    caractéristiques articulatoires, 48

Julien Rouger
j.rouger@psy.gla.ac.uk
julien.rouger@gmail.com

Nationalité française
27 ans au 18 février 2008

Formation

2007-	Postdoctorat en neurocognition vocale - Voice Neurocognition Laboratory, Glasgow
2003-2007	Thèse de doctorat en neurosciences - Ecole doctorale CLESCO, Université Paul Sabatier, Toulouse
2001-2002	DEA acoustique mention bien - Ecole doctorale MEGA, Université Claude Bernard, Lyon
1999-2002	Etudes d'ingénieur généraliste (option acoustique en fin de cursus) - Ecole Centrale de Lyon
1996-1999	Etudes en classe préparatoire, spécialité Mathématiques Physique - Lycée Faidherbe, Lille
1996	Baccalauréat scientifique mention bien - Académie de Lille

Activités de recherche

• **À partir de décembre 2007** : "*Application of multivariate methods to functional Magnetic Resonance Imaging of brain activation patterns during human voice neurocognition*". Postdoctorat en neurocognition vocale, réalisé au Voice Neurocognition Laboratory (Department of Psychology, University of Glasgow). Partenariats : Department of Cognitive Neuroscience, University of Maastricht ; Department of Computer Science, University of Glasgow.

• **Octobre 2003 à novembre 2007** : "*Perception audiovisuelle de la parole chez le sourd postlingual implanté cochléaire et le sujet normo-entendant : étude longitudinale psychophysique et neurofonctionnelle*". Thèse de doctorat en neurosciences, réalisée au Centre Cerveau et Cognition UMR 5549 CNRS (Faculté de médecine de Toulouse-Rangueil) en partenariat avec le service ORL, le centre TEP et l'unité Inserm 825 du CHU Toulouse-Purpan.

• **Avril à septembre 2002** : "*Propagation d'un champ acoustique dans un conduit traité de révolution à section variable. Implémentation et validation du modèle de S. W. Rienstra*". Stage d'ingénieur fin d'études et de DEA acoustique, réalisé à Airbus France Toulouse, bureau d'études de Saint-Martin du Touch, département Acoustique et environnement.

• **Juin à août 2001** : "*Analyse astrophysique de photographies du ciel dans le proche infrarouge*". Stage d'ingénieur deuxième année, réalisé au Laboratoire d'astrophysique de l'observatoire Midi-Pyrénées UMR 5572 CNRS/INSU (Université Paul Sabatier, Toulouse).

Enseignement et encadrement pédagogique

• **Septembre 2006 à juin 2007** : prise en charge de cours de Biologie cellulaire et Neurocognition en 1re année de Psychomotricité à la Faculté de médecine de Toulouse-Rangueil.

• **Juin à juillet 2005** : encadrement d'un étudiant en stage de Master de médecine niveau I : "*Etude comportementale des interactions audiovisuelles chez le sujet sain et comparaison avec les performances des patients sourds implantés cochléaires*".

• **Septembre 2004 à juin 2005** : prise en charge de cours de Physiologie neuromusculaire en 1re année de Psychomotricité à la Faculté de médecine de Toulouse-Rangueil.

• **Septembre 2002 à octobre 2003** : prise en charge d'élèves de lycée en cours de soutien scolaire de Mathématiques et Physique pour la préparation aux examens du baccalauréat.

Publications scientifiques

✤ *Improvement in speech-reading ability by auditory training : Evidence from gender differences in normally hearing, deaf and cochlear implanted subjects.* Strelnikov K., **Rouger J.**, Lagleyre S., Fraysse B., Deguine O. & Barone P. (2008). Neuropsychologia, *in press.*

✤ *McGurk effects in deaf cochlear implant subjects.* **Rouger J.**, Deguine O., Fraysse B. & Barone P. (2008). Brain Res. 1188 :87–99.

✤ *Evidence that cochlear implanted deaf patients are better multisensory integrators.* **Rouger J.**, Lagleyre S., Fraysse B., Deneve S., Deguine O. & Barone P. (2007). Proc. Natl. Acad. Sci. U.S.A. 104(17) :7295–7300.

Communications scientifiques

* Janvier 2008 – 18es journées d'études de l'Association pour la promotion du langage parlé complété (ALPC) : *Implantation cochléaire : importance de la stimulation audiovisuelle pour optimiser la récupération de la parole* (**Rouger J.**)
* Novembre 2007 – 37th annual meeting of the Society for Neuroscience, San Diego, CA, USA : *Audiovisual integration of speech in postlingually deaf cochlear-implant patients : a behavioral and PET study* (Lagleyre S., **Rouger J.**, Démonet J.-F., Fraysse B., Deguine O., Barone P.)
* Octobre 2007 – 114e congrès national de la Société française d'ORL, Paris, France : *Reconnaissance de la voix chez le patient implanté cochléaire : étude comportementale comparative à un groupe de sujets normo-entendants* (Massida Z., Deguine O., **Rouger J.**, Lagleyre S., Laborde M.-L., Fraysse B., Belin P., Barone P.) ; *Meilleure intégration audiovisuelle de la parole chez les patients sourds postlinguaux implantés cochléaires* (**Rouger J.**, Lagleyre S., Fraysse B., Deneve S., Laborde M.-L., Deguine O., Barone P.) ; *Etude des mécanismes d'intégration audiovisuelle de la parole par l'effet McGurk chez les sujets sourds postlinguaux implantés cochléaires* (**Rouger J.**, Laborde M.-L., Deguine O., Fraysse B., Barone P.)
* Septembre 2007 – Autumn school From neural code to brain/machine interface, Montvillargenne, France : *Human cochlear implantation : neurofunctional and longitudinal aspects of audiovisual speech perception* (**Rouger J.**, Lagleyre S., Fraysse B., Démonet J.-F., Deguine O., Barone P.)
* Juin 2006 – 7th International multisensory research forum, Dublin, Ireland : *High visuo-auditory integration performances in deaf subjects with cochlear implants* (Barone P., Lagleyre S., **Rouger J.**, Laborde M.-L., Fraysse B., Deguine O.)
* Avril 2006 – Xe congrès de la Société internationale francophone d'ORL, Toulouse, France : *Apport et évolution de la bimodalité visuo-auditive dans la compréhension de la parole chez le patient implanté cochléaire* (Lagleyre S., **Rouger J.**, Laborde M.-L., Fraysse B., Barone P., Deguine O.)
* Octobre 2005 – 112e congrès national de la Société française d'ORL, Paris, France : *Etude par tomographie par émission de positons (TEP) des interactions visuo-auditives chez le patient implanté cochléaire* (Lagleyre S., **Rouger J.**, Laborde M.-L., Trotter Y., Barone P., Fraysse B., Deguine O.) ; *Apport de la bimodalité visuo-auditive dans la compréhension de la parole chez le patient implanté cochléaire* (Barone P., **Rouger J.**, Lagleyre S., Laborde M.-L., Trotter Y., Fraysse B., Deguine O.)
* Juin 2005 – 6th International multisensory research forum, Rovereto, Italy : *McGurk effects in deaf subjects with cochlear implants* (**Rouger J.**, Laborde M.-L., Lagleyre S., Trotter Y., Deguine O., Fraysse B., Barone P.) ; *Role of visuo-auditory integration in speech comprehension in deaf subjects with cochlear implants* (Lagleyre S., Laborde M.-L., **Rouger J.**, Trotter Y., Fraysse B., Barone P., Deguine O.)
* Août 2004 – Ecole d'été ACI Neurosciences computatives et intégrationnelles, Concarneau, France : *Réorganisation fonctionnelle chez le sourd implanté cochléaire* (**Rouger J.**, Barone P., Trotter Y., Deguine O.)

Compétences

Sciences appliquées

Sciences de la vie	Parole audiovisuelle, système auditif, implant cochléaire, neuroplasticité
Acoustique	Psychoacoustique, analyse et synthèse spectrale, traitement du signal, aéroacoustique
Mathématiques	Analyse et synthèse spectrale, traitement du signal, analyse matricielle, analyse psychophysique
Physique	Mécanique générale, mécanique des fluides, électromagnétisme, optique, physique nucléaire

Informatique et nouvelles technologies

Bureautique	Windows 98/2000/XP, Word, Excel, Powerpoint, Adobe Premiere
Langages	C++, Visual Basic, Matlab, LaTeX, Fortran 90, Pascal, Prolog, Maple, HTML, PHP, SQL
Algorithmique	Calcul numérique et matriciel, analyse et synthèse spectrale de sons et d'images

Langues étrangères

Anglais	Lu, parlé et écrit. Très bon niveau : GRE 520-800-660, TOEFL 620
Allemand	Lu, parlé et écrit. Niveau scolaire.
Espagnol	Notions élémentaires.

Audiovisual speech perception in postlingually deaf cochlear-implant patients and normally-hearing subjects : longitudinal psychophysical and neurofunctional study

Abstract

Our present work consisted in studying perceptual and neuronal mechanisms involved during audiovisual speech perception in postlingually deaf cochlear-implant patients and normally-hearing controls. With this aim in view we tested audiovisual performance for implanted patients during longitudinal follow-ups using behavioral methods and positron emission tomography neuroimagery, as well as incongruent audiovisual speech (McGurk paradigm). In order to achieve appropriate comparisons, control subjects were tested in identical conditions. Moreover they were tested in a speechreading condition and with auditory and audiovisual speech degraded through white-noise masks and cochlear-implant computer simulations. Results from these studies agree to point out, following cochlear implantation, a fast reorganization of multisensory cortical networks involved in audiovisual speech integration, leading to an optimal combination of audiovisual perceptual cues.

Keywords : audiovisual speech, cochlear implant, multisensory integration, cortical plasticity

Julien Rouger

Perception audiovisuelle de la parole chez le sourd postlingual implanté cochléaire et le sujet normo-entendant : étude longitudinale psychophysique et neurofonctionnelle

Thèse réalisée sous la direction de Pascal Barone et soutenue le 29 novembre 2007 à la Faculté de Médecine de Toulouse-Rangueil (Université Toulouse III — Paul Sabatier).

Résumé

Nos travaux ont consisté à étudier les mécanismes perceptifs et neuronaux impliqués lors de la perception audiovisuelle de la parole chez des patients sourds postlinguaux implantés cochléaires et des sujets contrôles normo-entendants. Dans ce but nous avons testé les performances audiovisuelles des patients implantés au cours de suivis longitudinaux en comportemental et en tomographie par émission de positrons, ainsi qu'au travers d'un paradigme de parole audiovisuelle incongruente (McGurk). Pour comparaison nous avons systématiquement testé des sujets contrôles dans les mêmes conditions. Nous avons également testé ces derniers en condition de lecture labiale puis en utilisant des dégradations de la parole auditive et audiovisuelle par du bruit blanc et par une simulation acoustique d'implant cochléaire. L'ensemble de ces études concorde pour indiquer, à la suite de l'implantation cochléaire, une réorganisation rapide des réseaux corticaux multisensoriels impliqués dans l'intégration audiovisuelle de la parole, aboutissant à une combinaison optimale des indices perceptifs audiovisuels.

Mots-clés : parole audiovisuelle, implant cochléaire, intégration multisensorielle, plasticité corticale

Discipline : Neurosciences

Laboratoire d'accueil :
Centre de recherche Cerveau et Cognition (CerCo)
UMR 5549 CNRS/UPS
Faculté de Médecine de Rangueil
31062 Toulouse Cedex 9