



HAL
open science

Adaptation des méthodologies d'évaluation sensorielle aux produits agroalimentaires à forte variabilité

Cécile Bavay

► **To cite this version:**

Cécile Bavay. Adaptation des méthodologies d'évaluation sensorielle aux produits agroalimentaires à forte variabilité. Alimentation et Nutrition. Université d'Angers, 2013. Français. NNT: . tel-00846841

HAL Id: tel-00846841

<https://theses.hal.science/tel-00846841>

Submitted on 22 Jul 2013

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.

Thèse de Doctorat

Cécile Bavay

Mémoire présenté en vue de l'obtention
du grade de Docteur de l'Université d'Angers
Sous le label de l'Université Nantes Angers Le Mans

Discipline : Science Alimentaire

Spécialité : Analyse sensorielle

Laboratoire : Unité de Recherche GRAPPE, SFR 4207 QUASAV
École Supérieure d'Agriculture d'Angers

Soutenue le 15 janvier 2013

École doctorale : ED VENAM (Végétal Environnement Nutrition Agroalimentaire
Mer)

Adaptation des méthodologies d'évaluation sensorielle aux produits agroalimentaires à forte variabilité

JURY

Rapporteurs : **Mme Sylvie Chollet**, HDR, Responsable du laboratoire d'analyse sensorielle,
Institut Supérieur d'Agriculture, Lille, France

M. Pascal Schlich, HDR, Directeur de Recherche, INRA, Dijon, France

Examineurs : **M. Per Bruun Brockhoff**, Professeur, Directeur de recherche, Technical
University of Denmark, Kongens Lyngby, Danemark

M. El Mostafa Qannari, Professeur, Oniris, Nantes, France

Directeur de Thèse : **Mme Emira Mehinagic**, HDR, Directrice de recherche, École Supérieure
d'Agriculture, Angers

Encadrants : **M. Ronan Symoneaux**, Ingénieur de recherche, École Supérieure d'Agriculture,
Angers

Millions saw the apple fall,
but Newton asked why.

Bernard Baruch
New York Post (24 June 1965)

Avant-Propos

Ce travail de recherche a été réalisé au sein de l'Unité de Recherche en Agroalimentaire sur les Produits et les Procédés (GRAPPE) du groupe ESA d'Angers, sous la direction scientifique d'Emira Mehinagic, responsable de l'Unité de Recherche GRAPPE et enseignant-chercheur à l'ESA, et le co-encadrement de Ronan Symoneaux, ingénieur de recherche à l'ESA, et Isabelle Maître, enseignant-chercheur à l'ESA. Ce projet, financé par la Région Pays de la Loire s'inscrit dans le cadre du programme « Nouvelle Équipe / Nouvelle Thématique », axé sur le développement des méthodologies sensorielles appliquées au matériel végétal.

Les résultats de ce travail de thèse ont fait l'objet des communications suivantes :

Publications

Bavay, C., Symoneaux, R., Maître, I., Kuznestova, A., Brockhoff, P.B., Mehinagic, E. (2013). Adaptation of sensory evaluation methodology for variable products: a study on apple. *Postharvest biology and technology*, 77, 67-74.

Bavay, C., Brockhoff, P.B., Maître, I., Kuznestova, A., Mehinagic, E., Symoneaux, R. (2013). Consideration of the sample heterogeneity and in-depth analysis of individual differences in sensory analysis. (*soumise le 22/10/2012 dans Food Quality and Preference*)

Kuznetsova, A., Christensen, R.H.B., **Bavay, C.**, Brockhoff, P.B. (2012). Automated Mixed ANOVA Modelling of sensory and consumer data. (*en préparation*)

Bavay, C., Mehinagic, E., Maître, I., Symoneaux, R. (2013). Towards adapted sensory methodologies for fresh produce evaluation: a review (*en préparation*)

Présentations dans des conférences internationales

Bavay, C., Symoneaux, R., Maître, I., Courcoux, P., Mehinagic, E. (2011). Non-destructive sampling to improve sensory descriptive results of varying products: an application on apple. In *9th Pangborn Sensory Science Symposium*, 4-8 September, Toronto, Canada (poster)

Bavay, C., Symoneaux, R., Maître, I., Mehinagic, E. (2012). Management of fruit presentation in sensory evaluation of apples for more reliable results. In *II International Symposium on Horticulture in Europe*, 1-5 July, Angers, France (poster)

Bavay, C., Symoneaux, R., Maître, I., Kuznestova, A., Brockhoff, P.B., Mehinagic, E. (2012). Methodological development for sensory evaluation of product presenting biological variability: a case study on apple. In *11th Sensometrics*, 10-13 July, Rennes, France (présentation orale)
Récompensée par le prix Addinsoft 'Best Young Researcher Award'

Bavay, C., Symoneaux, R., Maître, I., Mehinagic, E. (2012). Sensory profile adaptation to biological variability of fruit batch. In *5th European Conference on Sensory and Consumer Research*, 9-12 September, Bern, Switzerland (poster)

Bavay, C., Symoneaux, R., Maître, I., Kuznestova, A., Brockhoff, P.B., Mehinagic, E. (2012). The importance of methodological adaptation to get reliable results in sensory evaluation of product presenting biological variability: a case study on apple. In *SSP conference*, 10-12 October, Jersey City NJ, USA (poster)
Récompensée par le prix 'Elaine Skinner Award for Outstanding Presentation of an Applied Topic'

Remerciement

Je tiens tout d'abord à remercier les trois personnes qui m'ont encadrée (... et supportée) pendant ces trois années :

*Ma directrice de thèse, **Emira Mehinagic**, pour son encadrement. Merci pour tes conseils scientifiques et la réflexion apportée à mes travaux.*

Mes encadrants à la genèse de ce projet de thèse :

***Ronan Symoneaux**, pour sa confiance et sa bonne humeur. Merci pour ton aide. Je te remercie également d'avoir toujours été là. Quelle que soit mon humeur, tu as toujours su me faire rire.*

***Isabelle Maître**, pour son esprit critique. Merci de m'avoir aiguillonné sur les chemins de la recherche et de m'avoir soutenue.*

*Je remercie aussi les membres de mon comité de thèse, **Philippe Courcoux** et **François Laurens** pour leurs conseils avisés sur l'avancée de ce projet et pour leur sympathie.*

*Je remercie particulièrement **Per Brockhoff** et **Alexandra Kuznetsova** pour m'avoir accueillie à l'Université Technique du Danemark. Merci pour votre aide précieuse dans la compréhension des secrets de l'analyse de la variance... je ne les connais pas encore tous.*

*Je remercie **Sylvie Chollet** et **Pascal Schlich**, rapporteurs de ce mémoire, ainsi que **El Mostafa Qannari** et **Per Brockhoff** pour m'avoir fait l'honneur de juger ce travail de thèse.*

Je tiens ensuite à remercier toutes les personnes qui m'ont toujours aidée et soutenue dans la réalisation de ce travail :

***Isabel**, pour son aide précieuse dans la réalisation des séances de panel. Merci d'avoir coupé toutes ces pommes. Merci également pour ton investissement et ton soutien (tu as été un peu comme une maman).*

***Corinne**, pour la co-animation du panel. Merci d'avoir pris le relais quand j'en avais besoin. Merci également pour tous tes mots gentils.*

***Sophie**, pour son aide et sa disponibilité pour m'aider à apprivoiser la machine MTS. Merci pour tous ces petits trous effectués dans toutes ces pommes.*

***Dominique** (alias Rolie Madame Qualité), pour son aide dans les mesures instrumentales. Merci à toi aussi pour les petits trous.*

Je tiens particulièrement à remercier mes colocataires de bureau qui m'ont supportée et avec qui j'ai partagé beaucoup de moments agréables :

***Mathilde C.**, qui a été présente toute au long de ma thèse et avec qui j'ai partagé des moments particuliers. Tu as toujours été présente pour me soutenir et « me redonner la patate, patate, patate !! ». Grâce à toi j'ai fait pas mal de découvertes culinaires, toujours excellentes !*

***Stéphanie**, toujours souriante. Tu as su me détendre avec tes séances de tressage et me faire rire avec des vidéos toujours insolites !*

*Je tiens également à remercier notre voisine de bureau, **Céline**, qui a toujours eu une oreille attentive. Merci pour ton soutien et tous les moments partagés. Merci à toi et à Christopher de m'avoir fait l'honneur de ma présence éternelle dans votre jardin, j'espère que mes fruits seront aussi agréables que vous deux.*

*Je remercie aussi **Catherine**, pour ses conseils concernant la gestion de panel ; **Génica**, pour avoir partagé son expérience de thèse ; **Jérémy**, pour sa bonne humeur et ce surnom de « Boucle d'or » ; **Mara**, pour ses conseils précieux et les cours de Zumba partagés (ça y est, je réussis le pas de Samba !) ; **Mathilde D.**, pour le soleil qu'elle a dans le cœur ; **Laurence**, accueillante, souriante et qui a toujours répondu à toutes mes questions ; et **Aurélien** et **Véronique**, toujours disponibles et réactifs.*

Je remercie aussi tous les membres du GRAPPE pour les moments de détente autour d'un café, au détour d'un couloir ou lors des sorties de fin d'année.

*Non des moindres, je tiens à remercier, les membres du panel Eden, **Anita, Bernard, Brigitte B., Brigitte M., Corinne, Damaris, Florence G., Florence L., Joël, Karine, Laurent, Marianick, Marie-Geneviève, Marie-Josèphe, Martine, Patricia, Sandrine** et **Véronique**, sans qui cette thèse n'aurait pu avancer. Merci pour votre implication, votre motivation et votre convivialité, j'ai passé de super moments avec vous en dégustation et aux goûters.*

Je tiens aussi à remercier ma famille et mes amis pour leur soutien :

*Tout d'abord, **mes parents**, toujours présents et à l'écoute.*

*Ma cousine, **Gwen**, qui m'a fait passer des soirées inoubliables à chaque retour dans le Nord.*

*Mes amis ensuite, notamment **Milie, Malo, Ma Fillote, Raf**, qui m'ont soutenue, mais aussi **Papy, Awena** et **Pat** avec qui j'ai passé des soirées requinquantes.*

*Enfin, je remercie **Marco** qui a été d'un soutien indéfectible. Χίλια ευχαριστώ !*

Merci à tous de m'avoir supportée ... et supportée !!!

Sommaire

Introduction	1
Partie I. Revue bibliographique	4
1. La perception sensorielle.....	4
1.1 Mécanisme général de la perception sensorielle	4
1.1.1 La texture des pommes	5
1.1.1.1 Perception auditive	6
1.1.1.2 Perception somesthésique	6
1.1.2 Les saveurs des pommes (perception gustative).....	7
1.2 La perception sensorielle est influencée par différents facteurs	7
1.2.1 Les sujets d'un panel sont des individus uniques.....	8
1.2.2 Les conditions de dégustation conditionnent la qualité des résultats sensoriels..	9
1.3 Bilan de la perception sensorielle.....	10
2. Le profil conventionnel est un outil objectif de mesure des propriétés sensorielles des produits.....	11
2.1 Le profil a un rôle clé dans la mesure sensorielle	11
2.2 Le profil conventionnel permet d'obtenir une description complète des produits	11
2.3 Le panel entraîné est un outil de qualité	15
2.3.1 La qualité d'un outil est définie par la métrologie.....	15
2.3.1.1 La sensibilité ou discrimination	15
2.3.1.2 La fidélité ou répétabilité/reproductibilité.....	15
2.3.1.3 L'exactitude ou accord entre les sujets	16
2.3.1.4 La justesse ou concordance avec d'autres mesures.....	16
2.3.2 Analyse unidimensionnelle de la performance	17
2.3.2.1 L'analyse de la variance est très utilisée	17
2.3.2.2 Le facteur sujet : fixe ou aléatoire ?	19
2.3.2.3 Le terme d'interaction ne se réduit pas au simple désaccord	20
2.3.2.4 Adaptation du modèle standard à l'absence/la présence de répétitions....	22
2.4 Bilan sur l'outil d'évaluation sensorielle	23
3. La variabilité biologique des pommes dans la réponse sensorielle	24
3.1 La variabilité biologique est présente dans chaque lot	24
3.2 Les différences biologiques permettent de distinguer des variétés de pommes en fonction de leurs propriétés sensorielles	25
3.3 Les différences biologiques s'expriment aussi à l'échelle intra-variétale.....	26
3.4 Les différences biologiques s'expriment aussi à l'échelle intra-fruit	26
3.5 Certains facteurs externes accentuent la variabilité biologique	27
3.6 Bilan sur la variabilité biologique des propriétés sensorielles pommes.....	28
4. L'analyse et le traitement de la variabilité en évaluation sensorielle	29
4.1 L'analyse de la variance permet d'analyser les sources de variabilité de la réponse sensorielle	29
4.2 La correction des données pour éliminer les biais des sujets est courante	31
4.2.1 La correction de l'effet de niveau.....	31
4.2.2 La correction de l'effet dispersion	31

4.2.3	L'identification et l'élimination des sujets contribuant aux différences	32
4.3	Bilan sur l'analyse de la variabilité en évaluation sensorielle.....	33
5.	La maîtrise de la variabilité biologique en évaluation sensorielle	34
5.1	L'homogénéisation de l'échantillon pour réduire la variabilité biologique	34
5.1.1	L'échantillon analysé se doit d'être représentatif du lot.....	34
5.1.2	... mais en pratique, l'homogénéisation est courante	34
5.1.3	L'homogénéisation peut être réalisée sur des critères instrumentaux.....	35
5.1.3.1	Les mesures instrumentales peuvent être un indicateur de la qualité sensorielle.....	35
5.1.3.2	Mesure instrumentale destructive de la texture	36
5.1.3.3	Mesure instrumentale non destructive de la texture	37
5.2	La présentation des échantillons est un point clé en évaluation sensorielle	38
5.2.1	La présentation des échantillons répond souvent à des contraintes expérimentales	40
5.2.2	La présentation des échantillons est parfois utilisées pour réduire la variabilité	40
5.3	Bilan sur la maîtrise de la variabilité	41
6.	Conclusion	41
 Partie II. Problématique et démarche expérimentale		43
1.	Problématique et hypothèses	43
2.	Démarche expérimentale	45
2.1	Vérification de la sensibilité du panel	45
2.2	Adaptation du traitement des données sensorielles pommes.....	45
2.3	Réduction de la variabilité des fruits dégustés	46
 Partie III. Matériels et méthodes.....		48
1.	Notre produit modèle : la pomme	48
2.	La caractérisation sensorielle	49
2.1	Le panel entraîné	49
2.2	Les conditions de dégustation sont standardisées	49
3.	La mesure instrumentale de la texture	51
3.1	La fermeté mécanique, une mesure de référence	51
3.2	La fermeté acoustique, une mesure non-destructive	52
4.	Traitements statistiques	54
4.1	L'analyse de la variance	54
4.1.1	Test de significativité	54
4.1.2	Comparaison de moyenne.....	55
4.1.3	Estimation des composantes de la variance.....	56
4.2	La comparaison de variance	56
4.3	Corrélations.....	57
 Partie IV. Résultats et Discussion.....		58
 Chapitre I. Vérification de la sensibilité du panel		58
1.	Étude de la capacité de discrimination inter-variétale du panel.....	59
1.1	Sélection des fruits.....	59

1.2	Analyse de la variance : modèle standard	61
1.3	Résultats & discussion.....	62
1.4	Conclusion.....	65
2.	Étude la capacité de discrimination intra-variétale du panel.....	66
2.1	Sélection des fruits.....	66
2.2	Analyse de la variance : modèle mixte	67
2.3	Résultats & discussion.....	68
2.4	Conclusion.....	73
3.	Conclusion générale	73
Chapitre II. Mesurer la variabilité : sélection d'un modèle adapté		75
1.	Étude du modèle « fruit » incluant la prise en compte de la variabilité intra-lot	75
1.1	Les modèles.....	76
1.1.1	Le modèle standard	76
1.1.2	Le modèle « fruit » tenant compte de la variabilité intra-lot.....	76
1.1.3	Comparaison des résultats des deux modèles	78
1.2	Résultats et Discussion.....	79
1.2.1	Conséquence de l'ajout du facteur « fruit » dans le cas d'un jeu de données réelles	79
1.2.2	Impact d'une variabilité intra-lot très élevée (données simulées).....	84
1.3	Conclusion.....	86
2.	Étude de l'approche « assessor »	87
2.1	Observation des différences de notation des sujets	88
2.1.1	Démarche expérimentale	88
2.1.2	Sélection des fruits	89
2.1.3	Modèles mixtes.....	89
2.1.4	Les différences de notation des sujets	90
2.2	Approche « assessor »	95
2.2.1	Les modèles	95
2.2.1.1	Le modèle « fruit »	95
2.2.1.2	L'approche « assessor ».....	95
2.2.1.3	Comparaison des résultats des deux modèles	97
2.2.2	Résultats et Discussion	97
2.2.2.1	Conséquence de l'utilisation de l'approche « assessor » dans le cas d'un jeu de données réelles.....	97
2.2.2.2	Impact de la présence de dispersion seule ou combinée à du désaccord .	102
2.3	Conclusion.....	107
3.	Conclusion générale	108
Chapitre III. Diminuer la variabilité : échantillonnage et présentation des échantillons		109
1.	Étude de l'impact d'une homogénéisation instrumentale sur la variabilité de la réponse sensorielle	110
1.1	Méthode d'échantillonnage.....	110
1.2	Comparaison des échantillonnages	112
1.3	Résultats & discussion.....	113
1.3.1	Échantillonnage destructif.....	113
1.3.2	Échantillonnage non destructif.....	117

1.4	Conclusion.....	120
2.	Le partage des fruits par plusieurs sujets permet de réduire la variabilité sensorielle	122
2.1	Partage des fruits.....	122
2.2	Comparaison des méthodes de partage.....	122
2.3	Résultats et discussion.....	122
2.4	Conclusion.....	126
3.	Conclusion générale.....	129
Partie V. Conclusion et perspectives.....		130
Références.....		135
Index des tableaux.....		146
Index des illustrations.....		148
Annexes.....		150

Introduction

La qualité est définie comme l'aptitude d'un ensemble de caractéristiques intrinsèques à satisfaire des exigences (ISO, 2005c). La qualité des produits agroalimentaires se décline en qualité sanitaire et qualité alimentaire (Molnár, 1995). Cette dernière se décline elle-même en qualité sensorielle, qualité nutritionnelle et qualité d'usage. La qualité sensorielle est d'intérêt particulier puisqu'elle est déterminante dans l'acceptation d'un produit et permet également d'expliquer les préférences des consommateurs. C'est donc un outil d'aide à la prise de décisions stratégiques en industrie.

La qualité sensorielle est appréciée notamment par les méthodes descriptives d'analyse sensorielle. Ces méthodes ont été décrites par de nombreux auteurs comme simples, permettant d'obtenir le profil sensoriel complet des produits, et indispensables pour analyser les caractéristiques sensorielles et les intensités de ces caractéristiques responsables des préférences et des aversions (Stone & Sidel, 1993 ; Moskowitz 1993). Lors d'une analyse descriptive, les sujets doivent décrire les produits selon leurs caractéristiques sensorielles comme le ferait un instrument de mesure. L'épreuve la plus courante est le profil conventionnel. Cette méthode consiste à caractériser les produits selon une liste de descripteurs donnée. Elle est réalisée par un panel de sujets entraînés à l'évaluation des caractéristiques d'un type de produit. Chaque sujet note l'intensité de la perception de chaque descripteur. Pour un produit donné, les moyennes des réponses des sujets pour chaque descripteur constituent le profil sensoriel du produit.

Les réponses sensorielles présentent des variations pouvant être dues aux différences entre les sujets et/ou à l'hétérogénéité de l'échantillon proposé pour l'évaluation. D'une part, les différences entre les sujets constituent une source de variation inhérente aux différences génétiques ou acquises et au processus de perception. En effet, il existe des différences de perception et des différences

d'utilisation de l'échelle d'intensité entre les sujets. Les sujets diffèrent dans leur moyenne, leur dispersion sur l'échelle, leur répétabilité et leur classement des produits. L'entraînement du panel permet de réduire ces différences mais ne les élimine pas. D'autre part, les produits testés peuvent présenter des variations biologiques. Les produits végétaux illustrent très bien cette variabilité. Parmi eux, les fruits et notamment les pommes, ont été l'objet de très nombreuses études. La pomme est le fruit le plus cultivé au monde (69 569 612 tonnes en 2010). Elle est aussi le fruit le plus consommé en France (22.6% de la part de marché des fruits en volume), se plaçant devant l'orange (12.3%) et la banane (12.2%) (Frétillet, 2010). En outre, près de 90% des ménages Français achètent des pommes au moins une fois dans l'année (FranceAgrimer, 2012). Le choix de la pomme comme modèle pour de très nombreuses études n'est donc pas un hasard, particulièrement en Pays de la Loire, troisième région productrice de pomme (en volume) en 2010 (FranceAgrimer, 2012). La variabilité biologique des pommes a été soulignée à maintes reprises dans les différents travaux de développement de méthodes d'évaluation sensorielle des produits végétaux (Heintz et Kader, 1983, Stevens, M. A. et Albright, 1980, Williams, A. A. et Carter, 1977). Comme souligné par Hampson *et al.* (2000), la variabilité biologique des produits est un challenge car les différences au sein d'un échantillon peuvent rendre les différences entre les échantillons plus difficiles à détecter. La pomme a donc été choisie comme produit modèle pour chercher à adapter les méthodologies d'évaluation sensorielle aux produits agroalimentaires à forte variabilité.

L'objectif de ce travail de thèse est d'adapter les méthodologies sensorielles bien maîtrisées à la spécificité des produits subissant des variations qui impactent la qualité des lots étudiés, dont les produits végétaux font partie. Ce travail vise à :

- proposer des méthodes de référence fiables pour caractériser les produits soumis à une variabilité ;
- améliorer les performances des experts dans l'évaluation de produits soumis à une variabilité.

La mise en place de tels outils servira à une meilleure interprétation et exploitation des données issues de profils sensoriels grâce à l'identification des sources de la variabilité des données.

La première partie de ce manuscrit aborde la méthodologie d'évaluation sensorielle descriptive et présente l'état de l'art concernant la variabilité biologique de la pomme et des démarches entreprises pour y faire face.

La deuxième partie est consacrée à l'exposé de la problématique et de la démarche expérimentale.

La troisième partie présente les matériels et méthodes utilisés au cours du travail de thèse. L'exposé des résultats est organisé en trois chapitres. Le premier chapitre concerne l'étude de la sensibilité du panel. Il a pour but de vérifier la fiabilité de notre outil de mesure – le panel – et notamment sa capacité discriminante aux niveaux inter-lot et intra-lot, les lots correspondant à différentes variétés. Le deuxième chapitre présente l'adaptation du modèle d'analyse de la variance pour le traitement de données sensorielles issues de l'évaluation de produits variables. Les apports des modèles proposés sont évalués. Le troisième chapitre présente le travail effectué pour tenter de réduire la variabilité de la réponse sensorielle. Il a pour objectif d'évaluer l'efficacité de la réduction de la variabilité des échantillons sur la réduction de la variabilité de la réponse sensorielle.

Enfin, la conclusion propose des recommandations méthodologiques pour l'analyse sensorielle de produits pouvant présenter une variabilité biologique et l'exploitation des données.

Dans le souci de procurer un outil utile à la filière, nous avons concentré nos efforts de recherche sur l'utilisation des statistiques inférentielles. En effet, la filière a besoin d'outils permettant de tirer des conclusions sur les facteurs étudiés dans une population à partir des données provenant d'un échantillon de ses individus.

Partie I. Revue bibliographique

L'objectif de cette partie est, dans un premier temps, de dresser un état des lieux des pratiques d'évaluation sensorielle concernant la pomme. Suite à une description de la perception sensorielle, la méthode du profil conventionnel est décrite dans ses aspects sensoriels et sensométriques¹. L'importance de l'impact de la variabilité biologique des pommes sur la fiabilité des résultats issus de l'analyse sensorielle est ensuite exposée. Enfin, les différentes méthodes de traitement et de gestion de la variabilité en évaluation sensorielle et plus spécifiquement pour l'évaluation des pommes sont abordées.

1. La perception sensorielle

1.1 *Mécanisme général de la perception sensorielle*

Les perceptions sensorielles sont mesurées par un ensemble de méthodes regroupées sous le terme d'analyse sensorielle. Les stimuli sont de natures diverses, mais les étapes physiologiques de la perception sont toujours les mêmes (Figure 1). Lorsqu'un *stimulus* entre en contact avec les récepteurs sensoriels d'un être humain, il est converti en influx nerveux (transduction). Cet influx nerveux se propage jusqu'au système nerveux central (codage). Ce phénomène est appelé *sensation* (Depledge, 2009). L'information est d'abord traitée sans référence aux perceptions antérieures. Les

¹ La sensométrie regroupe l'ensemble des méthodes mathématiques et statistiques pour l'analyse des données sensorielles et de préférence.

résultats de ce traitement sont l'intensité du stimulus, correspondant à la masse globale d'activité nerveuse, et la qualité du stimulus, représentée par la répartition spatiale de l'activité. Lors de sa progression vers les centres supérieurs, l'information est filtrée, réduite et stabilisée. Dans les centres supérieurs, l'information est confrontée à la mémoire et à la conscience (intégration). Cette confrontation de la sensation avec la conscience correspond à la *perception*. Toute l'information est alors réunie dans un message global, sensoriel et affectif. L'être humain perçoit donc une image, une représentation du stimulus et non le stimulus lui-même (MacLeod *et al.*, 2009). La perception sensorielle est divisée en classes en fonction du système impliqué. Cinq classes sont définies : la perception visuelle, la perception olfactive, la perception gustative, la perception auditive et la perception somesthésique.

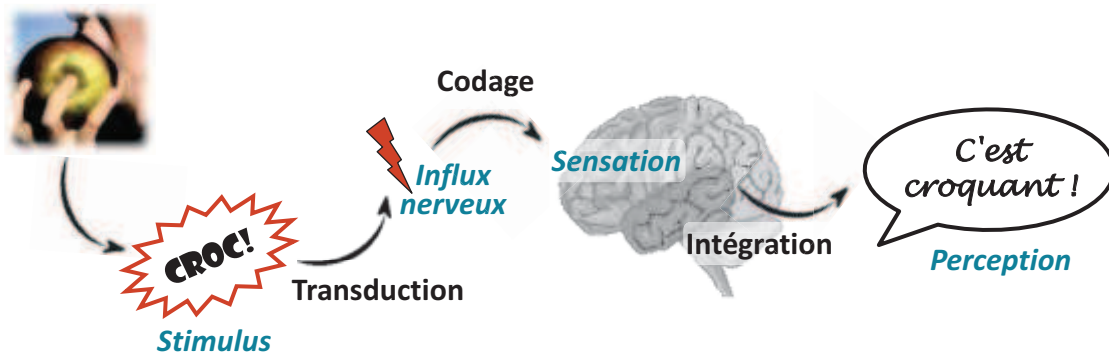


Figure 1. Schéma général de la perception sensorielle

Le travail de thèse s'est orienté uniquement sur la texture et la saveur de la pomme. C'est la raison pour laquelle seuls les mécanismes de perception impliqués dans leur évaluation sont mentionnés.

1.1.1 La texture des pommes

La texture est définie comme l'ensemble des propriétés mécaniques, géométriques et de surface d'un produit, perceptibles par les mécanorécepteurs, les récepteurs tactiles et, éventuellement, par les récepteurs visuels et auditifs (ISO, 1992). D'après Szczesniak (2002), cette définition véhicule des concepts importants tel le fait que la texture est une propriété sensorielle que seul l'Homme peut percevoir et décrire. Les instruments dits de mesure de la texture ne peuvent mesurer que certains paramètres physiques qui doivent ensuite être interprétés en termes de perception sensorielle. La texture est une propriété multidimensionnelle qui tient à la structure et à la composition de l'aliment. Ainsi, la texture des pommes est fonction des propriétés mécaniques des tissus, de leur densité, de la taille, la forme et l'adhérence des cellules, de l'épaisseur et de la résistance des parois cellulaires, ainsi que de la turgescence des vacuoles (Allan-Wojtas *et al.*, 2003, Jackman et Stanley,

1995). Elle dépend aussi de la composition chimique des cellules et notamment des parois cellulaires (Waldron *et al.*, 1997). Ces composantes peuvent intervenir dans la perception somesthésique et auditive de la texture que l'Homme traduit par des termes tels que : *dureté*, *croquant*, *jutosité*, *fermeté*, *masticabilité*, *granulosité* et *farinosité* (Daillant-Spinnler *et al.*, 1996, Harker *et al.*, 2002b, Mehinagic *et al.*, 2004).

1.1.1.1 Perception auditive

L'oreille est l'organe principal de l'ouïe. Les sons transmis par l'air sont captés et amplifiés par le pavillon qui les focalise vers le conduit auditif jusqu'au tympan, membrane qui entre alors en vibration. Ces vibrations sont transmises mécaniquement à l'oreille interne, par les osselets. Elles provoquent des ondes de pression correspondant aux ondes sonores. Ces ondes de pression permettent de communiquer les vibrations à l'organe récepteur, la cochlée. Les cils de l'oreille interne sont mis en mouvement, ce qui active la production d'influx nerveux chargés de transmettre l'information jusqu'au cortex auditif (Duizer, 2001). Un son est caractérisé par trois grandeurs. L'intensité, appelée aussi volume, permet de distinguer un son fort d'un son faible. La hauteur, liée à la fréquence, fait dire que ce son est plus ou moins aigu ou grave. Le timbre correspond à l'ensemble des harmoniques additionnées au son pur. Les descripteurs sensoriels des pommes *croquant* et *crissant* sont partiellement définis par l'intensité et la qualité du son produit respectivement lorsque le sujet croque la pomme et au cours de la mastication (Harker *et al.*, 2002b).

1.1.1.2 Perception somesthésique

Les sensations somesthésiques sont perçues par la peau, les muscles, les tendons et les articulations. Elles regroupent la sensibilité thermique, les sensibilités tactile et kinesthésique, résultant de contraintes mécaniques, mais aussi la sensibilité chimique résultant du contact direct de molécules avec les muqueuses (Depledt, 2009).

La sensibilité tactile fait intervenir la peau et les muqueuses, et permet d'évaluer des descripteurs sensoriels comme la *rugosité* de la peau, la *farinosité* de la chair et l'*astringence* (Mehinagic *et al.*, 2004). La sensibilité kinesthésique concerne les muscles, tendons et ligaments, et est impliquée dans la mesure de descripteurs tels que la *dureté*, la *fermeté*, la *masticabilité* et le *fondant* (Mehinagic *et al.*, 2004). Le descripteur *jutosité* fait intervenir les deux sensibilités tactile et kinesthésique.

1.1.2 Les saveurs des pommes (perception gustative)

La saveur est définie comme la sensation perçue par l'organe gustatif lorsqu'il est stimulé par certaines substances solubles (ISO, 1992). Ces substances sont des molécules chimiques en solution dans la salive. Les saveurs élémentaires sont les suivantes : acide, amer, salé, sucré, alcalin, umami et métallique (ISO, 1992). Pour la description des pommes les saveurs les plus utilisées sont la saveur acide et la saveur sucrée (Daillant-Spinnler *et al.*, 1996, Harker *et al.*, 2002c, Mehinagic *et al.*, 2004). On trouve parfois l'utilisation de la saveur amère (Daillant-Spinnler *et al.*, 1996).

Dans l'espace buccal, la sensation issue des stimuli passe par 7 à 10 millions de cellules sensorielles regroupées en bourgeons gustatifs et formant ce que l'on appelle communément les papilles gustatives. Elles sont particulièrement abondantes sur la langue mais se trouvent également sur les parois de la bouche, la voûte du palais, l'épiglotte et même sur le haut du pharynx. Contrairement à une idée longtemps admise, il n'y a pas de cellules vraiment spécialisées dans la reconnaissance d'une sensation ou d'une autre, même s'il existe des différences de sensibilité entre différents endroits de la langue (Lawless et Heymann, 1998c). La concentration en corps sapides conditionne la quantité de récepteurs activés et par conséquent l'intensité perçue de la saveur.

Les principaux corps sapides intervenant dans la perception des saveurs de la pomme sont le fructose, le glucose et le saccharose, responsables de la saveur sucrée et l'acide malique et l'acide citrique responsables de la saveur acide.

1.2 La perception sensorielle est influencée par différents facteurs

La perception sensorielle est le résultat d'un processus complexe qui intègre toutes les informations à disposition, provenant aussi bien du stimulus que de l'inconscient. De nombreuses études ont démontré l'impact de facteurs physiologiques et cognitifs tels que l'adaptation², l'interaction entre des stimuli (augmentation, suppression), et la fatigue physiologique. Par exemple, la texture, et donc la structure, peut influencer la perception des saveurs et des arômes (Brossard *et al.*, 2012). En effet, la déstructuration plus ou moins importante et plus ou moins rapide du produit en bouche impacte directement la libération des molécules sapides et volatiles. En parallèle, le relargage de molécules sapides comme les acides organiques peut provoquer une salivation plus importante et influence la perception de la texture (e.g. pour la jutosité). Des facteurs psychologiques sont également impliqués dans la construction de la réponse sensorielle : les idées préconçues dues aux informations

² Diminution ou changement de sensibilité à un stimulus suite à une exposition prolongée à ce stimulus ou un stimulus similaire

disponibles à propos de l'échantillon, l'accoutumance, l'association de deux propriétés d'un produit (e.g. les pommes vertes sont acides), le report lorsque deux propriétés sont liées (« halo ») ou lorsqu'un descripteur est manquant (« dumping »), l'ordre de présentation (Meilgaard *et al.*, 1991a). L'effet de « halo » correspond à l'augmentation de la note donnée à un descripteur à cause de la perception d'un autre descripteur (e.g. la perception d'un arôme citron peut augmenter la note donnée pour l'acidité d'un produit). L'effet de « dumping » correspond à l'augmentation de la note attribuée à un descripteur de par l'absence du descripteur permettant de décrire la sensation perçue par le sujet. Ces biais peuvent être minimisés en maîtrisant l'information donnée aux sujets, en contrôlant leur environnement de dégustation et en leur procurant un entraînement approprié (ISO, 1993). La réponse sensorielle peut également être influencée par des facteurs intrinsèques à l'individu comme son identité génétique (Bartoshuk, 2000), son âge (Bitnes *et al.*, 2007a) ou encore son expertise – entraînement à l'évaluation sensorielle et/ou connaissance des produits (Bitnes *et al.*, 2007a, Labbe *et al.*, 2004, Wolters et Allchurch, 1994).

1.2.1 Les sujets d'un panel sont des individus uniques

Chaque sujet est unique et présente des capacités uniques. On distingue trois types de capacités : les capacités liées à l'intellect comme la créativité et la mémoire, les capacités liées à la personnalité telles que l'aptitude à percevoir les réactions du corps, et enfin les capacités physiologiques aussi appelées seuils. Ces dernières influencent directement la perception et sont variables en fonction des individus (Blakeslee et Salmon, 1935).

La plupart des publications étudiant les capacités physiologiques des sujets se sont focalisées sur la mesure de la perception des composés sapides et odorants. Par exemple, concernant les saveurs, Bartoshuk *et al.* (1994) ont étudié l'influence de l'anatomie sur les seuils et ont rapporté une sensibilité plus élevée chez les sujets possédant un plus grand nombre de bourgeons gustatifs.

Les études de seuils concernant la texture sont rares. Cela s'explique par la complexité de sa perception : la texture est multidimensionnelle et implique plusieurs types de récepteurs (Szczesniak, 2002). Une étude a mené à l'établissement d'un seuil : comparant les mesures sensorielles et instrumentales de fermeté, Harker *et al.* (2002b) ont observé qu'une différence minimale de 6N était nécessaire pour qu'une différence texturale (*croquant*, *crissant* et *jutosité*) soit perçue par des sujets entraînés.

Ces différences individuelles impliquent des variations dans la réponse de différents sujets à un même stimulus. Dans le cadre d'une analyse descriptive, une mesure de la tendance centrale, le plus souvent la moyenne arithmétique, est utilisée pour décrire la réponse à un stimulus. Cette mesure est appropriée si la distribution des scores du groupe est symétrique et unimodale. Dans le cas

contraire, l'utilisation de cette moyenne peut produire des résultats inexacts. Par ailleurs, la taille des variations des données en fonction des stimuli est jugée par rapport à la taille des variations intra-stimulus. Ainsi, la probabilité de détecter une différence entre des stimuli diminue lorsque les variations entre les sujets augmentent. Les seuils sont donc parfois abordés comme un problème de variation indésirable (Stevens, D. A., 1996). Deux approches de ce problème ont été proposées : éliminer la source de variation ou incorporer la variation dans l'analyse. L'élimination totale de cette variation est utopique. Il a cependant été montré qu'il est possible d'améliorer les capacités sensorielles avec l'entraînement et l'exposition (Bitnes *et al.*, 2007a, Pangborn, 1959).

1.2.2 Les conditions de dégustation conditionnent la qualité des résultats sensoriels

La presque totalité des facteurs cités comme influençant la perception sensorielle sont liés aux conditions de dégustation choisies par l'expérimentateur. Le respect des bonnes pratiques permet une maîtrise des conditions liées à l'environnement de la dégustation (température constante et confortable, bonne ventilation et air inodore, bruits parasites limités, sujets isolés). La présentation des échantillons est également importante. Le mode de présentation et le nombre d'échantillons sont deux paramètres cruciaux (Piggott, 1995). Pour l'évaluation des pommes, l'usage est d'effectuer les dégustations en lumière rouge (Mehinagic *et al.*, 2004, Seppä *et al.*, 2012) ou de peler les échantillons (Barreiro *et al.*, 1998, Harker *et al.*, 2002b) afin de minimiser les idées préconçues ou les associations cognitives liées à la couleur des échantillons. Le nombre d'échantillons à présenter au cours d'une séance doit être choisi judicieusement afin de limiter la fatigue sensorielle (Moskowitz, 1983). Par ailleurs, afin de limiter l'impact des effets d'adaptation et lorsque le nombre de produits à évaluer est supérieur ou égal à trois, il est recommandé de choisir un plan de présentation équilibré (MacFie *et al.*, 1989). Le carré latin est couramment utilisé et consiste en un carré dans lequel chaque ligne et chaque colonne comporte tous les traitements. Les carrés latins de Williams équilibrent l'effet de position et l'effet de report entre produits (Williams, E.J., 1949). Chaque produit est présenté un même nombre de fois en première, deuxième... et dernière positions. Et, chaque produit est présenté un même nombre de fois avant chaque autre. Les carrés latins mutuellement orthogonaux (MOLS) permettent aussi une maîtrise de l'ordre de présentation des échantillons, avec chaque produit présenté un même nombre de fois en première, seconde... et dernière position. De plus avec cette méthode, chaque produit est dégusté avant chaque autre produit le même nombre de fois et deux crans avant chaque autre produit (Wakeling et MacFie, 1995).

1.3 *Bilan de la perception sensorielle*

Le passage de la sensation à la perception sensorielle est un processus complexe, et les facteurs influençant la perception sensorielle sont nombreux et ne sont pas toujours maîtrisables. Des variations dans la réponse à un stimulus sont donc observées en fonction des sujets, des moments d'évaluation, des conditions d'évaluations, etc. Ainsi, la mesure sensorielle est parfois qualifiée de « méthode subjective ». Cependant, une méthode objective est définie comme « une méthode dans laquelle les effets de l'influence personnelle sont minimisés » (Sauvageot, 1982). Dans ce cadre, l'évaluation sensorielle peut être considérée comme une méthode objective puisque les bonnes pratiques tendent à réduire l'influence des facteurs externes liés aux sujets et aux conditions de dégustation. De plus, des méthodologies d'évaluation sensorielle adaptées aux objectifs de discrimination ou de description de produits ont été développées. Parmi celles-ci, le profil conventionnel permet de décrire de la façon la plus exhaustive et la plus objective possible les caractéristiques des produits d'un groupe.

2. Le profil conventionnel est un outil objectif de mesure des propriétés sensorielles des produits

2.1 *Le profil a un rôle clé dans la mesure sensorielle*

Deux types d'épreuves sensorielles existent. Elles répondent à deux objectifs distincts : discriminer et décrire. Les épreuves descriptives sont plus complexes et également plus riches. Elles permettent d'obtenir des informations sur les propriétés caractéristiques des produits. Les profils sensoriels obtenus sont une information clé dans le suivi de la qualité, le développement de produits et l'étude des préférences consommateurs. Stone et Sidel (2004) considèrent cette méthode comme indispensable pour déterminer les profils complets des produits et analyser les caractéristiques et les intensités des caractéristiques influençant les choix et l'appréciation. MacLeod *et al.* (2009) et Moskowitz (1983) soulignent que le profil sensoriel permet d'obtenir une image sensorielle des produits complète et éloquente.

2.2 *Le profil conventionnel permet d'obtenir une description complète des produits*

Les méthodes de profil descriptif sont nombreuses. Certaines méthodes consistent en la description consensuelle de certaines propriétés : les méthodes Flavor profile® (Cairncross et Sjostrom, 1950) et Texture profile® (Brandt *et al.*, 1963). D'autres méthodes permettent une description complète des produits en intégrant des données individuelles par sujet : la méthode Spectrum® développée par Civile (Meilgaard *et al.*, 1991b), la méthode Quantitative Descriptive Analysis (QDA®) (Stone et Sidel, 1998, Stone *et al.*, 1974), le profil libre (Williams, A.A et Arnold, 1984) ou encore le profil flash (Dairou et Sieffermann, 2002). La méthode du profil conventionnel (ISO, 2003) est une généralisation de la méthode QDA®.

Le profil conventionnel a pour but d'effectuer une description complète des propriétés sensorielles d'un ensemble de produits sur le plan qualitatif et quantitatif. Les produits constituent un univers et sont évalués en relatif, selon une liste de descripteurs préalable. Le profil conventionnel est réalisé par un groupe de sujets sélectionnés et entraînés. La synthèse de l'information des sujets permet de constituer les profils sensoriels des produits. Les moyennes des réponses des sujets pour chaque descripteur constituent ces profils sensoriels.

La méthodologie est basée sur la recherche de descripteurs permettant de donner un maximum d'informations sur les propriétés sensorielles des produits étudiés puis sur la quantification de ces descripteurs. La réalisation des profils passe par plusieurs étapes (ISO, 1994) (Figure 2). La première étape consiste en la sélection d'un panel. Les participants sont recrutés en fonction de leur acuité sensorielle, leur capacité d'expression face aux produits et leur motivation. La motivation est d'ailleurs une qualité essentielle pour assurer la dynamique du groupe (Murray *et al.*, 2001). Il existe différents tests qui peuvent être utilisés afin de vérifier les aptitudes sensorielles et non sensorielles des candidats (ISO, 1993). Les tests de sélection ne semblent pas prédire de façon sûre les performances des sujets en profil (Lesschaeve et Issanchou, 1996). Il convient cependant de s'assurer des capacités des sujets en fonction de celles qui seront mises en œuvre dans les futures études sensorielles (Lawless et Heymann, 1998a). Une fois le panel constitué, il convient d'expliquer au panel les principes généraux de l'analyse sensorielle et de l'entraîner aux méthodes d'évaluation. Lorsque le panel est prêt, l'espace produit sélectionné peut lui être présenté afin qu'il génère un maximum de termes permettant de décrire l'ensemble des produits. Le choix de l'espace produit conditionne les descripteurs qui seront générés par le panel et est donc important. La liste de termes ainsi obtenue est réduite en supprimant les termes hédoniques, les synonymes, les antonymes et les termes non pertinents. Il est aussi d'usage de choisir les termes appartenant à la terminologie existante. Le consensus et le tri statistique peuvent être utilisés dans cette procédure. Les définitions verbales et/ou les références physiques, et le protocole d'évaluation de chaque descripteur, ainsi que l'ordre d'évaluation des descripteurs sont déterminés par consensus. Par exemple, pour l'évaluation des pommes, les saveurs utilisées sont principalement l'*acide* et le *sucre* (Daillant-Spinnler *et al.*, 1996, Harker *et al.*, 2002c, Mehinagic *et al.*, 2004). Les descripteurs de texture de la pomme sont présentés dans Le Tableau 1. Les termes utilisés diffèrent parfois mais les définitions se rejoignent. L'importance du vocabulaire et des définitions associées est ici mise en exergue. Les sujets sont ensuite entraînés à l'utilisation de cette liste de descripteurs, le but étant de mettre les sujets d'accord sur la sensation associée à chaque descripteur et d'aligner leurs concepts sensoriels. Il est possible de contrôler les performances des sujets et du panel pendant cette phase d'entraînement afin d'orienter le travail à effectuer en cours de l'apprentissage, mais aussi, à la fin de l'entraînement afin de valider la qualité du panel (Lawless et Heymann, 1998a).

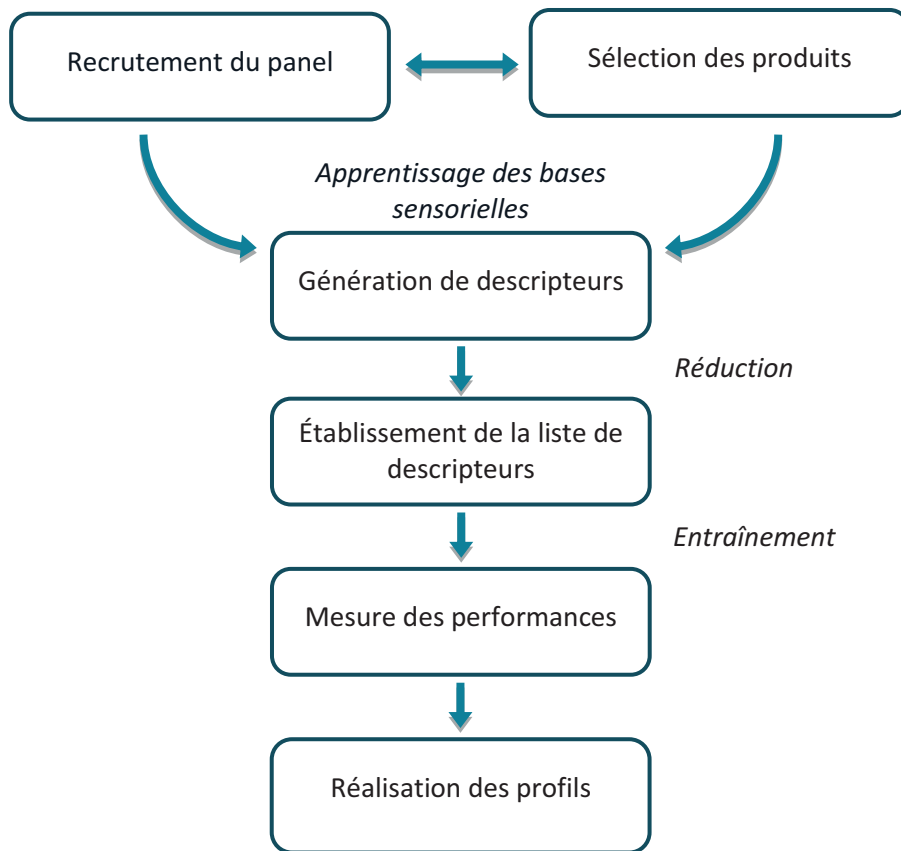


Figure 2. Étapes pour l'établissement d'un profil sensoriel (ISO, 1994)

Tableau 1. Principaux descripteurs de la texture des pommes

Attributs	Définition	Référence
<i>Dureté de la peau '1st bite – Skin toughness'</i>	Mesure de la dureté de la peau	(Dailliant-Spinnler <i>et al.</i> , 1996)
<i>Jutosité '1st bite – Juiciness'</i>	Quantité de jus libérée lorsqu'on mord pour la 1 ^{ère} fois avec les incisives	
<i>Dureté '1st bite – Hardness'</i>	Résistance à l'action d'enfoncer les incisives dans l'échantillon	
<i>Crissant 'During chewing – Crispness'</i>	Degré du son entendu au cours de la mastication	
<i>Jutosité ' During chewing – Juiciness'</i>	Quantité de jus libérée au cours de la mastication	
<i>Fondant ' During chewing – Slow breakdown'</i>	Temps nécessaire pour dégrader la chair et pouvoir l'avaler (sans la peau)	
<i>Spongieux ' During chewing – Spongy'</i>	Texture ressemblant à éponge	
<i>Croquant 'Crispness'</i>	Son généré lorsqu'on mord pour la 1 ^{ère} fois avec les incisives	(Harker <i>et al.</i> , 2002b)
<i>Fermeté 'Firmness'</i>	Force nécessaire pour enfoncer les incisives dans l'échantillon	
<i>Jutosité initiale 'Initial Juiciness'</i>	Quantité de jus libérée de l'échantillon pendant les trois 1ères mastications avec les molaires	
<i>Crissant 'Crunchiness'</i>	Son généré lors des mastications avec les molaires	
<i>Facilité de dégradation 'Ease of breakdown'</i>	Nombre de mastications nécessaire pour dégrader la chair et pouvoir l'avaler	
<i>Jutosité 'Sustained juiciness'</i>	Quantité de jus libérée de l'échantillon pendant une mastication prolongée	
<i>Pulpeux 'Pulpiness'</i>	Quantité de matière humide et fibreuse qui apparaît pendant la mastication	
<i>Granuleux 'Mealiness'</i>	Capacité de la chair à devenir des petites particules	
<i>Farineux 'Flouriness'</i>	Capacité de la chair à devenir de très petites particules sèches	
<i>Résistance au toucher 'Touch resistance'</i>	Résistance du fruit à la pression exercée par la main	
<i>Rugosité 'Roughness'</i>	Degré de rugosité de la peau mesurée par le toucher	
<i>Croquant 'Crunchiness'</i>	Force nécessaire pour arracher un morceau de pomme et évaluation du son qui en résulte	
<i>Masticabilité 'Chewiness'</i>	Durée et nombre de mastications (à compter) nécessaires pour mâcher la pomme (chair et peau) afin de la rendre prête à être avalée	
<i>Jutosité 'Juiciness'</i>	Quantité de jus libéré par la pomme au cours de la mastication	
<i>Farinosité 'Mealiness'</i>	Texture friable et sèche qui tend vers celle de la pomme de terre cuite	
<i>Fondant 'Fondant'</i>	Force nécessaire pour désagréger un morceau de pomme moyennement mastiqué placé entre le palais et la langue	

2.3 Le panel entraîné est un outil de qualité

2.3.1 La qualité d'un outil est définie par la métrologie

L'analyse sensorielle se doit de refléter au plus juste les propriétés sensorielles des produits étudiés. En métrologie, les qualités d'un outil de mesure sont sa sensibilité, sa fidélité, son exactitude et sa justesse (ISO, 2007b). Ces critères constituent la fiabilité. Ils sont applicables à l'outil d'évaluation sensorielle, dont la qualité est définie par la performance, généralement définie elle-même par trois critères : discrimination, accord et répétabilité (ISO, 1993).

2.3.1.1 La sensibilité ou discrimination

La *sensibilité* est le quotient de la variation d'une indication d'un système de mesure par la variation correspondante de la valeur de la grandeur mesurée (ISO, 2007b). La sensibilité est liée à la plus petite variation sur l'entrée (stimulus) qui puisse causer un changement visible dans l'indication de la mesure. En analyse sensorielle, la sensibilité correspond à la *discrimination* qui est la capacité de la méthode à différencier les produits évalués, lorsque des différences existent. Il est important de noter que le pouvoir discriminant est relié à la répétabilité mais aussi à l'accord entre les sujets (Pineau, 2006). Plus le niveau de répétabilité et/ou l'accord entre les sujets est/sont faible(s), plus la discrimination sera difficile. A l'inverse, un pouvoir discriminant élevé indique que les sujets sont suffisamment répétables et suffisamment en accord pour qu'un consensus sur les différences entre les produits soit obtenu. L'utilisation de ce critère pour valider une méthode peut être mise en cause puisqu'il ne s'agit pas en réalité d'obtenir le plus de différence possible entre les produits mais de refléter la réalité des différences perçues.

2.3.1.2 La fidélité ou répétabilité/reproductibilité

La *fidélité* est l'étroitesse de l'accord entre les indications ou les valeurs mesurées obtenues par des mesures répétées du même objet ou d'objets similaires dans des conditions spécifiées (ISO, 2007b). La répétabilité et la reproductibilité peuvent être vérifiées.

La *répétabilité* est définie comme l'étroitesse de l'accord entre les résultats successifs obtenus avec la même méthode sur une matière identique soumise à l'essai dans les mêmes conditions (même opérateur, même appareil, même laboratoire et court intervalle de temps) (ISO, 2005b). En évaluation sensorielle, la répétabilité est donc la capacité à donner une même note à un même stimulus évalué au cours d'une même séance (Etaio *et al.*, 2010). La répétabilité peut être évaluée au

niveau individuel ou au niveau du groupe. La répétabilité du panel peut être considérée comme un critère de validation. Il est en effet important qu'une mesure répétée par un même panel sur les mêmes stimuli et dans un intervalle de temps court fournisse des résultats permettant d'aboutir aux mêmes conclusions.

La *reproductibilité* est définie comme l'étroitesse de l'accord entre les résultats successifs obtenus avec la même méthode sur une matière identique soumise à l'essai en faisant varier les conditions (ISO, 2005b) (changement de lieu ou de moment, par exemple). En évaluation sensorielle, la notion de reproductibilité n'est pas définie de la même manière par tous. Pour Rossi (Rossi, 2001), la reproductibilité correspond à l'étroitesse de l'accord entre un sujet et le reste du panel. Cependant, cette notion d'accord entre les sujets est rarement considérée comme critère de reproductibilité. La reproductibilité peut correspondre à la capacité de différents panels à donner une même note à un même stimulus (Barylko-Pikielna *et al.*, 1992), ou à la capacité d'un panel à donner une même note à un même stimulus évalué au cours de séances différentes (Etaio *et al.*, 2010, Evin et Siekierski, 2002).

2.3.1.3 L'exactitude ou accord entre les sujets

L'*exactitude* est l'étroitesse de l'accord entre une valeur mesurée et une valeur vraie d'un mesurand³ (ISO, 2007b). L'exactitude de mesure est quelquefois interprétée comme l'étroitesse de l'accord entre les valeurs mesurées qui sont attribuées au mesurand. En évaluation sensorielle, l'exactitude peut alors être interprétée comme *l'accord entre les sujets*. Ce critère est la conséquence immédiate de la particularité de l'outil sensoriel qui se compose de plusieurs unités contrairement aux autres outils de mesure « classiques ». Le niveau d'accord entre les sujets est évalué pour connaître la variabilité autour de la mesure.

2.3.1.4 La justesse ou concordance avec d'autres mesures

La *justesse* est l'étroitesse de l'accord entre la moyenne d'un nombre infini de valeurs mesurées répétées et une valeur de référence (ISO, 2007b). Il s'agit souvent d'établir la correspondance entre les valeurs données par la méthode à valider avec des valeurs obtenues indépendamment par un autre instrument correctement calibré (Lawless et Heymann, 1998b). La mesure sensorielle doit par exemple refléter les opinions des consommateurs, ou peut être comparée à des mesures instrumentales. Toutefois, dans la plupart des cas, la complexité des produits ne permet pas une telle approche (Pineau, 2006) et il est difficile d'établir une référence « vraie ».

³ Grandeur que l'on veut mesurer

2.3.2 Analyse unidimensionnelle de la performance

Les méthodes de contrôle de performances disponibles sont nombreuses. Elles couvrent les champs unidimensionnels (étude d'un seul descripteur) (Labbe *et al.*, 2004, Næs, 1998, Næs *et al.*, 2010, Pagès et Husson, 2009, Rossi, 2001, Schlich, 1994, 1997, Tomic *et al.*, 2007) et multidimensionnels (étude simultanée de plusieurs descripteurs) (Bitnes *et al.*, 2007b, Couronne, 1997, Dahl, Tobias et Næs, 2004, Dahl, T. et Næs, 2009, Dijksterhuis, 1995, Findlay *et al.*, 2006, Husson *et al.*, 2005, Huynh *et al.*, 2012, King, M. C. *et al.*, 2001, Monrozier et Danzart, 2001, Sinesio *et al.*, 1991, Thybo et Martens, 2000).

Parmi toutes ces méthodes, l'analyse de la variance est la plus couramment utilisée. La suite de ce paragraphe se focalise sur cette méthode.

2.3.2.1 L'analyse de la variance est très utilisée

Le principal objectif des méthodes d'analyse de la variance est d'identifier les facteurs responsables de la variabilité de la réponse et de quantifier la part de responsabilité de chacun de ces facteurs. En pratique, cela est le plus souvent réalisé en identifiant d'abord quels sont les facteurs qui sont importants, puis en étudiant les facteurs les plus importants plus en détail par l'utilisation de tests post-hoc. Le modèle le plus couramment utilisé en évaluation sensorielle, et que nous appellerons *modèle standard*, est le suivant :

$$X_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijr} \quad (1-1)$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet}^* \text{variété}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ représente l'effet moyen global, ν_i correspond à l'effet fixe produit, α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction (aléatoire) sujet*produit et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Ce modèle est applicable aux données à « vraies » répétitions, c'est-à-dire lorsque les répétitions ne présentent pas une structure définie du type séance ou lot de fabrication (Næs *et al.*, 2010, Schlich, 1997). L'indice r indique les répétitions. L'analyse standard passe par la définition des sommes des carrés des écarts (SCE), des carrés moyens (CM) et des ratios de carrés moyens (statistique de Fisher F). Les calculs de ces paramètres ainsi que ceux des degrés de liberté (ddl) sont détaillés dans le Tableau 2.

Tableau 2. Paramètres du modèle d'ANOVA standard

Effet	Ddl	SCE	CM	F
Produit	I-1	$JK \sum_i (x_{i..} - x_{...})^2$	$\frac{SCE_{produit}}{ddl_{produit}}$	$\frac{CM_{produit}}{CM_{sujet*produit}}$
Sujet	J-1	$IK \sum_j (x_{.j.} - x_{...})^2$	$\frac{SCE_{sujet}}{ddl_{sujet}}$	$\frac{CM_{sujet}}{CM_{sujet*produit}}$
Sujet*Produit (interaction)	(I-1)(J-1)	$K \sum_i \sum_j (x_{ij.} - x_{i..} - x_{.j.} + x_{...})^2$	$\frac{SCE_{sujet*produit}}{ddl_{sujet*produit}}$	$\frac{CM_{sujet*produit}}{CM_{erreur}}$
Erreur	IJ(R-1)	$\sum_i \sum_j \sum_r (x_{ijr} - x_{ij.})^2$	$\frac{SCE_{erreur}}{ddl_{erreur}}$	

Un effet sujet significatif indique que tous les sujets n'ont pas le même niveau moyen de notation. Par exemple, en Figure 3, le sujet 2 présente un niveau de notation moyen supérieur au sujet 1. L'entraînement ne permet pas totalement de corriger les différences de niveau entre les sujets. Cet effet est souvent significatif mais n'apporte pas d'information sur les performances du panel. L'effet produit correspond à la discrimination. Un effet produit significatif indique que les notes moyennes des produits sont statistiquement différentes et le panel est déclaré discriminant pour le descripteur étudié. La statistique de Fisher de l'effet produit et la probabilité associée sont les indicateurs les plus couramment utilisés pour juger la discrimination d'un panel. Un effet de l'interaction sujet*produit significatif informe que les sujets ne perçoivent pas les mêmes différences entre les produits. L'interaction est alors communément utilisée comme un indicateur de l'accord des sujets. Un effet de l'interaction significatif est donc interprété comme un désaccord entre les sujets. Cependant, le désaccord (exemple en Figure 3) n'est pas la seule composante expliquant la différence de perception des produits qui est reflétée dans ce terme d'interaction. Les différences d'étendue des notes, c'est-à-dire de dispersion sur l'échelle sont également exprimées dans l'interaction (Brockhoff, 2003). Dans l'exemple de la Figure 3, le sujet 2 utilise la totalité de l'échelle de notation alors que le sujet 1 tend à donner des notes rapprochées aux trois produits. Enfin, le terme d'erreur exprime la différence entre les notes des répétitions de chaque produit. Il peut être regardé comme un indicateur de la répétabilité du panel : une erreur faible est synonyme de bonne répétabilité du panel.

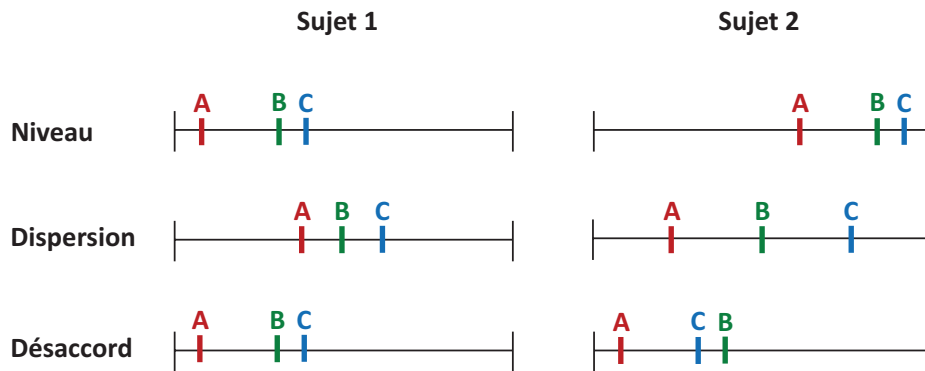


Figure 3. Les différences d'utilisation de l'échelle entre les sujets

2.3.2.2 Le facteur sujet : fixe ou aléatoire ?

Le choix du facteur sujet en effet fixe ou aléatoire a longtemps fait débat. Lundahl et McDaniel (1988) ont été parmi les premiers à discuter ce thème. Ils démontrent la pertinence de l'utilisation du facteur sujet en effet aléatoire. Dans le cas de panels de consommateurs, il ne fait aucun doute que les sujets ont été tirés aléatoirement parmi une population. Par contre, dans le cas de panels entraînés, le débat est orienté vers la sélection et l'entraînement des sujets. Les partisans du sujet en effet fixe avancent que la sélection et l'entraînement rendent le tirage des sujets non aléatoire. Cependant, l'entraînement ne change rien à la population initiale dans laquelle ont été tirés les sujets et des performances similaires sont attendues de tout nouveau sujet introduit dans le panel. Par ailleurs, la mesure de la reproductibilité des résultats (au sens de la comparaison de panel) implique de considérer le facteur sujet en effet aléatoire.

Les auteurs discutent également les conséquences du choix du facteur sujet en effet fixe ou aléatoire en termes statistiques. Lorsque le sujet est en effet fixe, sous l'hypothèse de nullité de l'effet produit, le carré moyen attendu de l'effet produit est égal au carré moyen de l'erreur. La source de variation produit est donc testée contre l'erreur résiduelle.

$$F_{produit} = \frac{CM_{produit}}{CM_{erreur}} \quad (I-2)$$

Lorsque le sujet est en effet aléatoire (comme dans le modèle (I-1)), sous l'hypothèse de nullité de l'effet produit, le carré moyen attendu de l'effet produit est égal au carré moyen attendu de l'interaction sujet*produit. Les différences en termes de carrés moyens attendus en fonction du choix de l'effet sujet fixe ou aléatoire sont exposées dans le Tableau 3.

La source de variation produit est donc testée contre l'interaction.

$$F_{produit} = \frac{CM_{produit}}{CM_{sujet*produit}} \quad (1-3)$$

Si l'interaction est faible, les résultats sont presque inchangés. En revanche, lorsque l'interaction est forte, les conclusions peuvent être différentes. De plus, lorsque le sujet est en effet aléatoire, le nombre de degrés de liberté du dénominateur est plus petit, les différences entre les produits doivent donc être supérieures pour obtenir une statistique de Fisher significative.

Tableau 3. Carrés moyens attendus selon la nature de l'effet sujet dans le modèle à deux facteurs produit et sujet et leur interaction

Effet	CM attendu avec facteur sujet en effet fixe	CM attendu avec facteur sujet en effet aléatoire
Produit	$JK \sigma_{produit}^2 + \sigma_e^2$	$JK \sigma_{produit}^2 + K \sigma_{sujet*produit}^2 + \sigma_e^2$
Sujet	$IK \sigma_{sujet}^2 + \sigma_e^2$	$IK \sigma_{sujet}^2 + K \sigma_{sujet*produit}^2 + \sigma_e^2$
Sujet*Produit (interaction)	$K \sigma_{sujet*produit}^2 + \sigma_e^2$	$K \sigma_{sujet*produit}^2 + \sigma_e^2$
Erreur	σ_e^2	σ_e^2

Enfin, le choix du facteur sujet en effet fixe implique que les résultats obtenus ne s'appliquent qu'aux sujets et aux produits de l'étude (Næs et Langsrud, 1998). Le but de l'analyse sensorielle étant généralement d'obtenir une caractérisation des produits généralisable à toute la population, l'utilisation du facteur sujet en effet aléatoire est préférable.

Pour conclure, l'introduction du facteur sujet en effet aléatoire permet de généraliser les conclusions à la population, de vérifier la reproductibilité des résultats et d'éviter des conclusions erronées à propos des différences entre les produits.

2.3.2.3 Le terme d'interaction ne se réduit pas au simple désaccord

Le terme d'interaction peut être décomposé comme suit (Couronne, 1997) :

$$\begin{aligned} SCE_{interaction} &= K \sum_i \sum_j (x_{ij} - x_{i..} - x_{.j.} + x_{...})^2 \\ &= K \sum_i \sum_j ((x_{ij} - x_{.j.}) - (x_{i..} - x_{...}))^2 \\ &= K \sum_i \left(\sum_j d^2(\text{sujet } j \text{ centré}, \text{sujet moyen centré}) \right) \end{aligned} \quad (1-4)$$

L'interaction correspond à la somme des carrés des distances entre la note moyenne de chaque sujet et la note moyenne du panel, après centrage des données par sujet. Différents cas de figure peuvent mener à la même estimation de l'interaction. La Figure 4 illustre les cas d'une interaction due au désaccord entre les deux sujets (à gauche) et aux différences de dispersion sur l'échelle (à droite). Dans ce dernier cas, le classement des produits est le même pour les deux sujets.

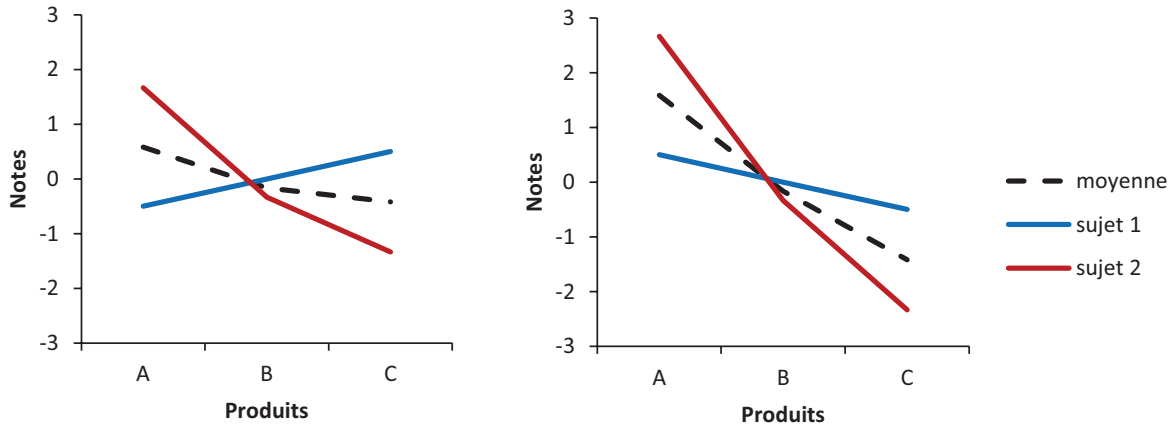


Figure 4. Le terme d'interaction : représentation graphique de ses deux interprétations

Fort de ce constat, Brockhoff et Skovgaard (1994) ont proposé un modèle, baptisé « assessor model », qui intègre cette dispersion en un effet :

$$X_{ijr} = \beta_j \cdot v_i + \alpha_j + \varepsilon_{ijr} \quad (I-5)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_j^2) \text{ indépendants entre eux}$$

Où v_i correspond à l'effet fixe produit, α_j à l'effet aléatoire sujet, β_j au coefficient de dispersion du sujet j et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle. Ce modèle autorise l'hétérogénéité des variances des notes des sujets, contrairement au modèle standard. Il peut être combiné à d'autres modèles afin de tester les performances des sujets (Brockhoff, 2003). Par exemple, la comparaison du modèle (I-5) au modèle (I-1) permet de tester la présence de désaccords entre les sujets. Une version améliorée de ce modèle est décrite par Brockhoff *et al.* (2012). Le modèle proposé comprend un terme d'interaction, reflet du désaccord, et utilise les notes moyennes centrées des produits comme co-variables :

$$X_{ijr} = \mu + v_i + \alpha_j + m_i \cdot \beta_j + \gamma_{ij} + \varepsilon_{ijr} \quad (I-6)$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet} \cdot \text{produit}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où ν_i correspond à l'effet fixe produit, α_j à l'effet aléatoire sujet, β_j au coefficient de dispersion du sujet j , m_i à la note moyenne centrée du produit i , γ_{ij} à l'interaction sujet*produit (désaccord) et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

2.3.2.4 Adaptation du modèle standard à l'absence/la présence de répétitions

Lorsque les répétitions des mesures correspondent à différentes séances d'évaluation (données à « fausses » répétitions), un effet séance peut être ajouté au modèle. Les interactions avec ce facteur peuvent aussi être ajoutées (Schlich, 1994). Le modèle devient :

$$X_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \delta_r + \omega_{ir} + \lambda_{jr} + \varepsilon_{ijr} \quad (1-7)$$

$$\alpha_j \sim N(0, \sigma_{sujet}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{sujet*produit}^2)$$

$$\delta_r \sim N(0, \sigma_{séance}^2)$$

$$\omega_{ir} \sim N(0, \sigma_{séance*produit}^2)$$

$$\lambda_{jr} \sim N(0, \sigma_{séance*sujet}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où ν_i correspond à l'effet fixe produit, α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction sujet*produit, δ_r à l'effet séance, ω_{ir} à l'interaction séance*produit, λ_{jr} à l'interaction séance*sujet et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Le calcul des effets des interactions est effectué avec le carré moyen de l'erreur. Le calcul des effets simples nécessite la combinaison de plusieurs carrés moyens. Par exemple, le calcul de la statistique F pour l'effet produit devient :

$$F_{produit} = \frac{CM_{produit}}{CM_{sujet*produit} + CM_{séance*produit} - CM_{erreur}} \quad (1-8)$$

Le calcul des degré de liberté est ici effectué en utilisant l'approximation de Satterthwaite (Satterthwaite, 1946).

À l'inverse, lorsque les données ne présentent pas de répétitions, le modèle (I-1) est réduit de l'interaction :

$$X_{ij} = \mu + \nu_i + \alpha_j + \varepsilon_{ij} \quad (I-9)$$

$$\varepsilon_{ij} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où ν_i correspond à l'effet fixe produit, α_j à l'effet aléatoire sujet et ε_{ij} à l'erreur résiduelle (aléatoire) du modèle.

Les statistiques F des effets produit et sujet sont alors calculées avec le carré moyen de l'erreur.

Nous reviendrons sur les modèles d'analyse de la variance en lien avec l'explication de la variabilité de la réponse sensorielle dans la Partie I.4.1.

2.4 Bilan sur l'outil d'évaluation sensorielle

La structure des données de profil se prête bien aux analyses statistiques. Le profil conventionnel est un outil métrologique efficace. En raison de son système d'évaluation monadique séquentiel, le profil conventionnel permet l'évaluation d'un grand nombre de produits sur plusieurs séances. De plus selon MacLeod et Strigler (2009), le profil conventionnel est le meilleur outil disponible pour tenter d'expliquer les choix du consommateur. La question de la qualité des résultats sensoriels et de la nécessité de contrôler l'accord entre les sujets et son impact sur la réponse sensorielle a été traitée précocement, notamment par l'analyse de la variance et son modèle standard. Cependant, il semble que l'évaluation des pommes pose un problème additionnel, lié à la nature variable de ce produit. Par exemple, Williams et Carter (1977), dans leur méthodologie d'évaluation des pommes Cox's Orange Pippin, souligne la complexité de l'interprétation des résultats sensoriels pomme due à la confusion des sources de variabilité sujets et hétérogénéité du produit. Le paragraphe suivant est consacré à l'exposé de cette variabilité biologique qui est inhérente à la pomme et à tous les produits du végétal spécialisé.

3. La variabilité biologique des pommes dans la réponse sensorielle

Les auteurs ayant contribué au développement de l'évaluation sensorielle des produits végétaux, dont la pomme fait partie, ont constaté que l'évaluation de ces produits posaient certains problèmes spécifiques tels que la non disponibilité de référence lors de la phase d'entraînement, l'évolution des produits dans le temps et l'hétérogénéité des produits, aussi appelée variabilité biologique (Heintz et Kader, 1983). L'interprétation des résultats sensoriels peut alors s'avérer compliquée. Les fortes variations des propriétés sensorielles entre les individus d'un même lot peuvent induire une difficulté à déterminer la source des différences de notation : des différences dans la compréhension des descripteurs, des différences de sensibilité des sujets ou encore les différences effectives entre les individus peuvent être mises en cause. Hampson *et al.* (2000) insistent sur les conséquences de la variabilité biologique et avancent que les différences effectives au sein d'un lot donné peuvent rendre les différences entre les lots plus difficiles à détecter. Pour ces auteurs, la variabilité biologique est un « challenge ». D'autres auteurs, utilisateurs de l'évaluation sensorielle (e.g. Mehinagic *et al.* (2003)), avancent l'hypothèse de l'hétérogénéité des fruits pour expliquer les différences de notation des sujets.

La pomme, fruit d'intérêt économique, a fait l'objet de nombreuses études visant à identifier les sources de variabilité de ses propriétés, dans le but de comprendre les variations observées et de maîtriser les variations indésirables. Ce paragraphe expose les variations effectives pouvant exister entre les fruits d'un lot et contribuer à la variabilité des résultats sensoriels à différentes échelles.

Avant de continuer, il nous paraît nécessaire de clarifier la notion de lot et le vocabulaire utilisé dans la suite du manuscrit. Un *lot* est un ensemble d'individu produits ou fabriqués dans des conditions présumées uniformes (AFNOR, 1991). Dans notre étude, les lots correspondent à différentes variétés de pommes. Un *individu* est un objet concret, autrement-dit une pomme dans notre étude. Un *échantillon* est défini comme un ensemble d'individus prélevés au hasard dans le lot.

3.1 La variabilité biologique est présente dans chaque lot

La variabilité biologique liée à la nature du fruit est présente entre les différentes variétés de pommes (e.g. Watada et Abbot (1985)), entre les fruits d'une même variété cultivés dans des vergers différents (e.g. Róth *et al.*(2007)) mais aussi dans un verger unique (e.g. Broom *et al.* (1998)), voire entre les fruits provenant d'un arbre unique (e.g. Broom *et al.* (1998)) (Figure 5). Par ailleurs, des

variations sont observées au sein d'un fruit unique, en fonction de la position sur le fruit et de l'orientation de la dégustation (e.g. Dever *et al.* (1995)) (Figure 5). Cette variabilité est présente dans chaque lot de pommes et peut être large, puisqu'un lot est constitué de plusieurs pommes, provenant de plusieurs arbres, appartenant potentiellement à différents vergers. Elle est plus ou moins importante en fonction de la variété. Il semble alors que certains génotypes aient tendance à présenter des niveaux de maturité différents en fonction des fruits et à subir les conséquences du stockage.

Les paragraphes suivants décrivent les différences observées aux différents niveaux indiqués en Figure 5.

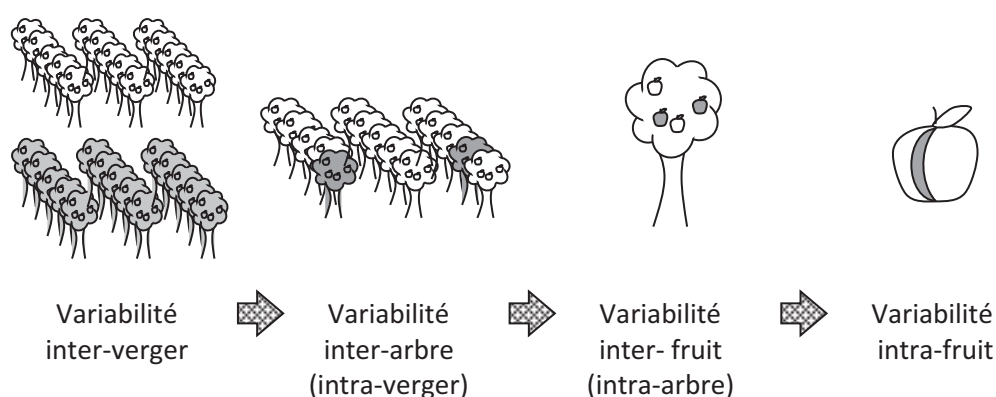


Figure 5. La variabilité biologique s'exprime à différents niveaux

3.2 Les différences biologiques permettent de distinguer des variétés de pommes en fonction de leurs propriétés sensorielles

Les variations des propriétés sensorielles des pommes sont d'abord observées entre des variétés différentes. De nombreux auteurs ont étudié les propriétés sensorielles des pommes et démontré des différences inter-variétales (Brookfield *et al.*, 2011, Feliciano *et al.*, 2010, Hampson *et al.*, 2000, Jaeger *et al.*, 1998, Karlsen *et al.*, 1999, Kühn et Thybo, 2001, Mehinagic *et al.*, 2003, Watada *et al.*, 1980). Par exemple, Watada *et al.* (1980) ont montré que les variétés Golden Delicious, Miller Spur, Redspur, York Imperial et Rome Beauty ne présentaient pas les mêmes propriétés sensorielles. Par exemple, la variété Rome Beauty était évaluée moins juteuse et plus farineuse que les quatre autres variétés. Par ailleurs, King *et al.* (2000) ont étudié les propriétés de clones de deux variétés de pomme et ont démontré que différents clones d'une même variété présentaient des différences pour certains descripteurs sensoriels.

3.3 Les différences biologiques s'expriment aussi à l'échelle intra-variétale

Mehinagic *et al.* (2003) ont émis l'hypothèse que les différences de notation observées entre les sujets étaient dues à l'hétérogénéité des fruits puisque chaque sujet dégustait une pomme différente. De même, des travaux visant à démontrer l'efficacité d'un petit groupe de dégustateurs pour évaluer la qualité des pommes ont suggéré la variabilité intra-lot comme cause possible du désaccord entre les sujets (Brookfield *et al.*, 2011), notamment en début de saison quand le développement des propriétés des fruits est rapide.

Les différences biologiques sont en étroit lien avec l'hétérogénéité physiologique des fruits liée à la maturité (Peirs *et al.*, 2001, Tijssens *et al.*, 2010). Parmi les études sensorielles, Watada *et al.* (1980) ont caractérisé plusieurs variétés de pommes récoltées à quatre dates : deux semaines et une semaine avant la date de maturité optimale, à maturité optimale et une semaine après. Les résultats montrent une décroissance générale des notes d'intensité des propriétés texturales des pommes et de l'acidité et une augmentation des notes pour les descripteurs *farineux* et *sucré* avec la maturité. L'importance de ces changements est dépendante de la variété. Notamment, les pommes de la variété Golden Delicious présentent peu de changement de texture, avec par exemple des notes allant de 2,3 à 3,1 pour la *dureté* et de 1,7 à 2,6 pour la *farinosité*. Au contraire, pour les pommes de la variété Redspur, les changements texturaux sont conséquents avec la *dureté* qui passe de 4,5 à 2,8 et la *farinosité* de 0,8 à 2,3 (sur échelle allant de 0 à 7). D'autres études plus récentes confirment ces résultats avec l'observation de notes plus élevées pour les descripteurs de texture tels que *fermeté*, *croquant* et *jutosité* dans le cas des modalités basses de maturité des pommes (Harker *et al.*, 2002b, Mehinagic, 2004, Mehinagic *et al.*, 2006, Seppä *et al.*, 2012). Par ailleurs Seppä *et al.* (2012) ont également démontré que l'influence de la maturité dépend de la variété de pomme étudiée.

3.4 Les différences biologiques s'expriment aussi à l'échelle intra-fruit

Nous avons vu que la variabilité biologique s'exprime entre les fruits de différentes variétés mais aussi entre les fruits d'une même variété. Mais, la variabilité biologique s'exprime aussi au sein d'un fruit. Les produits végétaux sont souvent de nature anisotropique. L'anisotropie désigne la qualité d'un milieu dont les propriétés varient suivant la direction selon laquelle elles sont évaluées. Khan et Vincent (1993) ont testé l'impact de l'orientation (radial ou tangentiel) de la morsure des sujets sur la facilité à se briser de cubes de pomme. Les résultats ont montré que l'anisotropie des propriétés liées à la fracture de la chair de pomme pouvait aisément être perçue en bouche. Leur expérience

consistait notamment à interroger les sujets sur la facilité à se fendre des cubes de pommes et sur la force nécessaire pour enfoncer les dents dans le cube. Un fort consensus a été observé dans les affirmations que l'échantillon radial se fend plus facilement et que l'échantillon tangentiel nécessite plus de force. Certains auteurs se sont penchés sur les différences perceptibles entre les faces observables des pommes. Ces faces observées sur les pommes bicolores sont provoquées par l'ensoleillement qui induit une couleur rouge sur la face ensoleillée et une couleur de fond, verte ou jaune, sur la face non ensoleillée. Les sujets entraînés perçoivent la face ensoleillée des pommes comme plus sucrée que la face non ensoleillée (Dever *et al.*, 1995, Symoneaux *et al.*, 2002). Symoneaux *et al.* (2002) ont étudié les variétés Braeburn, Fuji et Idared et ont montré que la face ensoleillée des pommes était perçue plus ferme. Dever *et al.* (1995), quant à eux, ont travaillé avec les variétés McIntosh et Jonagold et ont montré que la face ensoleillée était perçue moins croquante. Ces derniers se sont aussi intéressés aux différences entre les parties de la pomme proches du pédoncule et opposées au pédoncule et ont montré que ces deux parties ont des propriétés sensorielles différentes (e.g. pour la variété Jonagold, la partie opposée au pédoncule est moins croquante). Quelle que soit les conclusions de ces études, un fait important et commun est mis en évidence : il existe des différences sensorielles perceptibles au sein de chaque fruit.

3.5 Certains facteurs externes accentuent la variabilité biologique

L'expression de ces différences est accentuée par les conditions de culture : nature des sols, climat, système de production, et les conditions de stockage : type et durée. Concernant les systèmes de production, des études ont comparé la qualité des produits issus de production conventionnelle et de production biologique. Weibel *et al.* (1998) ont montré une plus forte fermeté et des notes d'intensité des descripteurs de saveurs supérieures pour les pommes Golden Delicious issues de la culture biologique, comparées à celles issues de la culture conventionnelle. Un impact du mode de culture a été démontré sur la fermeté des pommes de variété McIntosh (DeEll et Prange, 1992). Les fruits issus du mode de culture biologique sont perçus plus fermes que ceux issus du mode de culture conventionnel.

Concernant le stockage, une diminution de la fermeté des fruits est généralement constatée (Jobling et McGlasson, 1995, Mehinagic *et al.*, 2004, Rizzolo *et al.*, 2010, Varela *et al.*, 2005, 2008, Watada *et al.*, 1980). Par exemple, les notes d'intensité des descripteurs *résistance au touché*, *croquant*, *masticabilité*, *jutosité* diminuent pendant le stockage, alors que les notes d'intensité des descripteurs *farineux* et *fondant* augmentent (Mehinagic *et al.*, 2004). Ces études ont également démontré le développement du descripteur aromatique *fruit mûr* (Varela *et al.*, 2008) et pour certaines variétés,

comme la Braeburn, l'augmentation des notes pour la saveur sucrée et la diminution des notes pour la saveur acide. Les résultats de l'influence de la durée de stockage varient en fonction des variétés et des descripteurs. Par exemple, Mehinagic *et al.* (2004) ont constaté que la variété Fuji ne montrait de changement significatif que pour la *jutosité*, alors que les variétés Golden Delicious et Braeburn montraient des changements pour ce même descripteur mais aussi pour les descripteurs de *fermeté au toucher* et *masticabilité*. Jobling et McGlasson (1995) ont conclu de la même manière que la variété Fuji conservait ses propriétés texturales au cours du stockage en atmosphère contrôlée, résultat associé à une faible production d'éthylène. L'influence des conditions de stockage a également été étudiée. Des différences ont été observées entre les fruits stockés sous atmosphère contrôlée ou sous atmosphère normale. Il a notamment été démontré que le stockage sous atmosphère contrôlée permettait de maintenir les propriétés sensorielles des pommes Gala (Saftner *et al.*, 2002) et des pommes Jonagored (Rizzolo *et al.*, 2010). Comparés aux fruits stockés en atmosphère normale, les fruits stockés en atmosphère contrôlée étaient plus fermes, croquants, acides et aromatiques, et moins farineux. Ce maintien s'explique notamment par la minimisation de la respiration des fruits et de la production d'éthylène.

3.6 Bilan sur la variabilité biologique des propriétés sensorielles pommes

La variabilité des propriétés sensorielles des pommes et les causes possibles de cette variabilité ont été largement étudiées. Cette variabilité apparait à différentes échelles : entre les fruits de différentes variétés, entre les fruits d'une même variété, au sein d'un même fruit ; et se retrouve dans chaque lot d'une variété donnée évalué par un panel entraîné (que la provenance des fruits soit unique ou multiple). Cette variabilité est d'abord d'origine biologique et est influencée par divers paramètres externes. Ces paramètres externes peuvent parfois être maîtrisés. La maîtrise de ces conditions externes est d'ailleurs conseillée par les auteurs de méthodologie sensorielle adaptée aux produits végétaux afin d'éliminer toutes variations dues à des effets non étudiés (Heintz et Kader, 1983). Les paragraphes suivant sont consacrés aux méthodes de mesure de la variabilité en évaluation sensorielle et à sa maîtrise.

4. L'analyse et le traitement de la variabilité en évaluation sensorielle

La variance représente une estimation de la dispersion des données par rapport à la moyenne, notée \bar{x} . Plus la variance est grande, plus les résultats sont étendus. Supposons N mesures d'un descripteur sensoriel, la variance est estimée comme suit :

$$s^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^2 \quad (I-10)$$

L'écart-type s est simplement la racine carrée de la variance s^2 . L'écart-type présente l'avantage d'être exprimé dans la même unité que les données.

Le degré de variabilité d'une mesure peut être évalué par ces indicateurs et par l'étendue inter-quartile (Meilgaard *et al.*, 1991b), et représenté par la densité de distribution de la mesure ou sous forme de boîte à moustaches. Ces outils graphiques permettent de visualiser facilement l'étendue des résultats ainsi que d'identifier des groupes de valeurs ou des valeurs aberrantes.

Concernant la réponse sensorielle, la variabilité observée est le reflet, d'une part, de la variabilité des sujets, et d'autre part, de la variabilité biologique des produits. La variabilité des sujets a rapidement été un sujet d'attention en évaluation sensorielle. Le besoin spécifique d'évaluer la fiabilité de l'outil 'panel' a conduit au développement de méthodologies de mesure de performance, de prise en compte de la variabilité des sujets et de correction de la réponse.

4.1 L'analyse de la variance permet d'analyser les sources de variabilité de la réponse sensorielle

Nous avons vu précédemment (Partie I.2.3.2.1) que la méthode la plus utilisée pour l'analyse des données de profil sensoriel est l'analyse de la variance (ANOVA).

Cette méthode est utilisée pour la mesure de performance et pour analyser les données d'une étude et conclure sur les différences entre produits. Elle permet de quantifier la part de chaque source de variabilité dans la variabilité globale de la réponse sensorielle. Le modèle standard (I-1) est couramment utilisé ainsi que les modèles (I-7) dans le cas de données avec répétitions présentant une structure particulière (séance ou lot identifiable) et (I-9) dans le cas de données sans répétition.

Dans certains cas, plusieurs facteurs pouvant influencer les propriétés sensorielles des produits sont étudiés simultanément et peuvent interagir. Dans une étude sur les effets de traitements améliorant la conservation de tranches de pommes, l'influence de la concentration en ascorbate de calcium, du temps de stockage et d'une atmosphère modifiée a été étudiée (Aguayo *et al.*, 2010). Le modèle

d'analyse de la variance utilisé pour tester l'influence de ces facteurs comprenait les trois facteurs cités ci-dessus ainsi que leurs interactions d'ordre deux. L'influence des trois facteurs a pu être démontrée ainsi que des interactions entre le type d'atmosphère et les facteurs concentration en ascorbate de calcium et temps de stockage. Par exemple, l'effet d'une atmosphère modifiée était plus prononcé pour les concentrations en ascorbate de calcium élevées.

Le modèle d'analyse de la variance peut aussi être adapté à des conditions expérimentales particulières. Par exemple, Husson *et al.* (2004) ont étudié les différences sensorielles entre des cidres monovariétaux. Cinq variétés de pommes ont été sélectionnées et quatre produits ont été fabriqués pour chaque variété. Pour chaque variété, les quatre produits étaient fabriqués à l'identique. Dans le cas présent, le facteur produit n'est pas croisé avec le facteur variété mais imbriqué dans le facteur variété puisqu'un produit donné ne peut être lié qu'à une seule variété. Afin d'observer les différences entre les variétés, un modèle d'ANOVA mixte hiérarchique a été appliqué. Le modèle était le suivant :

$$X_{ijr} = \mu + A_i + \beta_j + A\beta_{ij} + C_{k(j)} + AC_{ik(j)} + \varepsilon_{ijk} \quad (I-11)$$

où μ est la moyenne générale, A_i est l'effet du sujet i , β_j est l'effet de la variété j , $A\beta_{ij}$ est l'interaction du sujet i avec la variété j , $C_{k(j)}$ est l'effet du produit k à l'intérieur de la variété j , $AC_{ik(j)}$ est l'interaction du sujet i et du produit k dans la variété j et ε_{ijk} est le terme d'erreur. Les effets β_j , $A\beta_{ij}$, $C_{k(j)}$, $AC_{ik(j)}$ et ε_{ijk} sont considérés comme indépendants et distribués normalement.

L'effet produit imbriqué dans l'effet variété a permis de prendre en compte les différences éventuelles entre les produits issus de la même variété de pomme.

L'analyse de la variance est une méthode très utilisée en évaluation sensorielle pour identifier les facteurs responsables de la variabilité des propriétés sensorielles. Le choix du modèle utilisé dépend des facteurs étudiés et des conditions expérimentales. Le choix du modèle est également influencé par la structure des données. Le choix du facteur sujet en effet fixe ou aléatoire a été discuté dans la Partie I.2.3.2.2. De la même manière chaque facteur du modèle peut être choisi fixe ou aléatoire. Cette décision dépend souvent de la manière dont l'échantillonnage est effectué et de la volonté d'extrapoler les résultats et a un impact sur la façon dont les effets des facteurs fixes sont testés (e.g. choix du facteur sujet en effet fixe ou aléatoire).

Quel que soit le modèle choisi, le facteur sujet est très souvent significatif. L'observation de la forte participation des sujets à la variabilité de la réponse sensorielle a mené au développement de méthodes de correction des données afin de limiter l'impact de cette variation indésirable.

4.2 La correction des données pour éliminer les biais des sujets est courante

4.2.1 La correction de l'effet de niveau

Le centrage de chaque sujet par rapport à la note nulle est courant. Ce prétraitement consiste à diminuer chaque note donnée par le sujet j de la moyenne du sujet j . Ainsi, tous les sujets ont un niveau de notation moyen égal à 0.

$$X'_{ijr} = X_{ijr} - X_{.j} \quad \text{avec} \quad X_{.j} = \frac{1}{IR} \sum_i \sum_r X_{ijr} \quad (I-12)$$

Où i est l'indice produit, j l'indice sujet et r l'indice répétition.

Lorsque les données sont traitées par ANOVA, l'effet de niveau est pris en compte lorsqu'un effet sujet est introduit dans le modèle.

4.2.2 La correction de l'effet dispersion

Plusieurs stratégies ont été proposées pour tenir compte de l'effet de dispersion, consistant en une utilisation plus ou moins étendue de l'échelle. Le prétraitement le plus simple et le plus intuitif consiste à ramener les variances de chaque sujet à 1. Les notes de chaque sujet j sont divisées par l'écart-type de l'ensemble des notes du sujet j (Næs, 1990).

$$X'_{ijr} = \frac{X_{ijr}}{s_j} \quad \text{avec} \quad s_j = \sqrt{\frac{\sum_i \sum_r (X_{ijr} - X_{.j})^2}{IR - 1}} \quad (I-13)$$

Où i est l'indice produit, j l'indice sujet et r l'indice répétition.

La combinaison de la réduction et du centrage, décrit précédemment, est appelée standardisation.

Une autre méthode, plus complexe, consiste à minimiser les différences entre sujets afin de minimiser le critère suivant (Næs, 1990):

$$\text{critère} = \sum_j \sum_{j'} (c_j X_{ij} - c_{j'} X_{ij'})^2 \quad (I-14)$$

Les données corrigées sont obtenue en multipliant les notes originales par les constantes c_j estimées. Les constantes peuvent être très variables d'un sujet à l'autre, comme illustré dans Næs (1990). Certains sujets présentent uniquement des constantes positives, synonyme de l'utilisation d'une plus petite partie d'échelle que le reste du panel. À l'inverse, les constantes de certains sujets sont toutes négatives. D'autres sujets varient fortement dans leur utilisation de l'échelle selon les descripteurs et leurs constantes varient alors fortement d'un descripteur à l'autre.

Une autre méthode a été proposée par Romano *et al.* (2008). Cette technique consiste à ajuster l'« assessor model » (Brockhoff et Skovgaard, 1994) aux données, puis à calculer la différences entre les données brutes et les valeurs prédites par le modèle (erreurs résiduelles).

Les différentes méthodes ont été comparés (Romano *et al.*, 2008) en mettant l'accent sur la capacité à réduire les interactions dans les tests d'ANOVA. Il a été constaté que tous les procédés donnaient les mêmes résultats. Romano *et al.* (2008) indiquent que l'utilisation de la standardisation n'est pas justifiée lorsque les variances résiduelles ne sont pas proportionnelles entre elles et que les tailles des désaccords individuels, mesurés par la variance des contributions au désaccord pour chaque sujet, ne sont pas proportionnels à l'utilisation individuelle de l'échelle.

Plus récemment, Brockhoff a proposé une autre transformation des données suivant l'équation suivante (Pineau, 2006):

$$X'_{ijr} = X_{ijr} - m_i \cdot (\beta_j - \bar{\beta}) \quad (I-15)$$

Où m_i est la moyenne des notes du produit i , β_j est le coefficient de dispersion du sujet j et $\bar{\beta}$ est le coefficient de dispersion du panel.

Cette transformation permet de faire coïncider la dispersion du sujet j avec la dispersion du panel. Pineau (2006) a appliqué cette transformation dans le cadre d'une étude du modèle « assessor » et a ouvert la discussion sur la pertinence de cette transformation lorsque l'accord entre les sujets n'est pas parfait.

4.2.3 L'identification et l'élimination des sujets contribuant aux différences

Des méthodes d'évaluation de la contribution de chaque sujet à l'interaction (Schlich, 1994) ou utilisant les tests de contrastes pour évaluer l'écart de chaque sujet à la moyenne du panel (Lundahl, D. S. et McDaniel, 1990) ont été proposées pour identifier les sujets participants au désaccord. Les auteurs observent ensuite l'évolution du terme d'interaction lorsque des sujets sont éliminés. L'élimination de sujets *a posteriori* est aventureuse et discutable puisqu'elle va à l'encontre du principe du profil qui est de caractériser des produits par l'avis du groupe. Par ailleurs, l'expérimentateur est alors en position de juge en décrétant la perception de certains sujets fausse. De plus dans le cas de produits présentant une forte variabilité biologique, l'erreur de jugement attribué au sujet pourrait se révéler être due aux propriétés intrinsèques de l'individu testé par ce sujet.

4.3 Bilan sur l'analyse de la variabilité en évaluation sensorielle

Le développement de méthodes d'analyse spécifiques pour l'étude des données sensorielles a notamment conduit à l'adaptation des modèles d'analyse de la variance pour la mesure non seulement des variations dues aux différences entre les produits étudiés mais aussi des variations entre les sujets du panel. La source de variabilité que sont les sujets d'un panel a largement été étudiée et différentes méthodes de prise en compte de cette variabilité ont été proposées. Cependant, les méthodes proposées, comme le modèle « assessor » restent peu utilisées par les analystes sensoriels. La Partie I.3 faisait état de la variabilité biologique présente dans les lots de pommes soumis à l'évaluation sensorielle. L'objectif de l'analyse de la variance, qui est d'identifier les facteurs responsables de la variabilité de la réponse et de quantifier leur part de responsabilité, n'est donc pas totalement rempli selon nous. Pour ce faire, il nous semble que la variabilité biologique (ou variabilité intra-lot) devrait être prise en compte. La variabilité intra-lot a été abordée dans les études de la qualité des pommes, mais pas prise en compte dans les analyses de données. Le traitement de la variabilité intra-lot concerne le plus souvent une volonté de réduire cette variabilité. La partie suivante décrit les méthodes employées en évaluation sensorielle des pommes pour réduire la variabilité intra-lot.

5. La maîtrise de la variabilité biologique en évaluation sensorielle

Même si la variabilité biologique est une caractéristique d'un produit, elle est souvent traitée comme une nuisance. L'expérimentateur a donc souvent recouru à des stratégies de réduction de la variabilité (Tijskens *et al.*, 2003).

5.1 *L'homogénéisation de l'échantillon pour réduire la variabilité biologique*

5.1.1 L'échantillon analysé se doit d'être représentatif du lot...

D'après les normes de contrôle des produits alimentaires (ISO, 1980), lors de l'analyse des caractéristiques d'un produit, l'échantillon prélevé dans le lot initial doit être représentatif du lot. Il est cependant spécifié qu'en présence de dommages dans le lot initial ou de forte hétérogénéité du lot initial, le lot doit alors être divisé en *sous-lots* homogènes qui seront analysés séparément et dans lesquels des individus seront prélevés au hasard. Pour les fruits et légumes, il est recommandé de prélever les individus au hasard dans le lot. La taille de l'échantillon dépend alors du calibre des produits : par exemple 3kg pour les pommes, qui sont des produits de taille moyenne (ISO, 1980).

5.1.2 ... mais en pratique, l'homogénéisation est courante

Les travaux de développement méthodologique pour l'évaluation de la qualité sensorielle des produits végétaux (Heintz et Kader, 1983, Stevens, M. A. et Albright, 1980) proposent d'homogénéiser les lots par les méthodes suivantes :

- Avoir des lots provenant de lieu de production avec des pratiques similaires ;
- Avoir des lots de niveaux de maturité similaires ;
- Limiter l'ampleur des changements physiologiques au cours de l'étude en minimisant la durée de celle-ci ;
- Choisir des individus sans défaut ni meurtrissure ;
- Couper les individus en cubes et les mélanger.

Ainsi la variabilité biologique est diminuée et interfère de façon réduite avec les objectifs de l'étude, ce qui permet de mettre en exergue les effets des facteurs étudiés.

5.1.3 L'homogénéisation peut être réalisée sur des critères instrumentaux

La pratique courante consiste à sélectionner des individus visuellement homogènes (taille, forme, couleur), sans défauts ni meurtrissures (Hampson *et al.*, 2000, Ioannides *et al.*, 2007, King, G. J. *et al.*, 2000, Konopacka et Plocharski, 2004); ou par souci d'objectivité à utiliser des mesures instrumentales pour sélectionner des fruits homogènes (Harker *et al.*, 2002a, Harker *et al.*, 2008, Konopacka et Plocharski, 2004).

5.1.3.1 Les mesures instrumentales peuvent être un indicateur de la qualité sensorielle

Les limites de l'évaluation sensorielle, notamment en termes de temps et de coût, ont conduit au développement de méthodes instrumentales de mesure des caractéristiques sensorielles. Les objectifs principaux visés par ces travaux sont de :

- Prédire les caractéristiques sensorielles ;
- Comprendre les caractéristiques sensorielles ;
- Expliquer les préférences des consommateurs.

Les mesures instrumentales utilisées pour mesurer les caractéristiques sensorielles des pommes peuvent être classées en deux catégories : destructives et non destructives.

Les descripteurs sensoriels de saveurs des pommes sont difficilement prédits par les mesures instrumentales. En effet, les corrélations observées entre les mesures biochimiques (concentration en composés sapides) et les mesures sensorielles sont faibles, excepté pour l'acidité perçue prédite par l'acidité titrable (Harker *et al.*, 2002c).

Concernant la texture des pommes, les méthodes proposées pour lier les paramètres sensoriels et la structure des produits sont nombreuses. Ces méthodes reposent sur diverses connaissances fondamentales, empiriques ou imitatives des mécanismes de rupture se produisant en bouche (Bourne, 2002). Les techniques de pénétrométrie et de compression sont les plus utilisées pour les mesures de texture. Par ailleurs, l'indice de Magness-Taylor, mesuré par pénétrométrie, a été utilisé efficacement pour classer des pommes au sein de variétés en fonction leurs propriétés texturales et a démontré un lien avec l'acceptabilité de la texture (Konopacka et Plocharski, 2004).

Les progrès techniques au cours des dernières décennies ont conduit à l'élaboration de dispositifs de contrôles non destructifs capables de mesurer des variables internes des fruits. Plusieurs types de méthodes non destructives de mesure de la texture des pommes ont vu le jour récemment et suivent divers principes (la réponse aux microdéformations, la réponse à un impact, la technique du rebond, la résonance magnétique nucléaire, les propriétés optiques et les propriétés vibrationnelles) (García-Ramos *et al.*, 2005). Parmi ces techniques non-destructives, les techniques acoustiques (Acoustic Firmness Sensor, AFS) et de microdéformations (Sinclair Internal Quality Firmness Tester, SIQ-FT) ont

été utilisées pour la classification de pommes (Harker *et al.*, 2008). Les résultats de l'étude consommateur qui a suivi ont démontré l'efficacité de ces méthodes dans l'échantillonnage des pommes en fonction de leur texture. Parmi ces deux méthodes, la mesure acoustique présente l'avantage d'être une mesure globale et donc de refléter les caractéristiques du fruit entier.

5.1.3.2 Mesure instrumentale destructive de la texture

Traditionnellement, la fermeté des fruits est estimée de manière destructive par la méthode de Magness-Taylor (pénétrométrie) (Harker *et al.*, 1996). Cette mesure est basée sur l'introduction d'un module cylindrique en inox dans la chair d'un fruit pelé pour mesurer la force de pénétration maximale. La mesure automatisée de pénétrométrie (Duprat *et al.*, 2000) permet de quantifier la résistance de l'épiderme et de la chair et donne accès à sept paramètres (force maximale de la courbe : F_s , déformation associée à F_s : D , pente ascendante : $Grad$, Force à 7mm : F_f , travail pour atteindre F_s et F_f : W_s et W_f respectivement, force limite de compression : FLC). Cette mesure est généralement bien corrélée aux mesures sensorielles puisqu'elle reproduit partiellement les effets de la mastication (Barreiro *et al.*, 1998, Harker *et al.*, 2002b, Mehinagic *et al.*, 2004). Mehinagic *et al.* (2004) ont mis en relation les paramètres extraits de la courbe résultant de la pénétration de la sonde dans le fruit et les descripteurs sensoriels de texture suivant : *masticabilité*, *jutosité*, *fondant*, *farinosité*, *croquant* et *résistance au touché*. Leurs résultats montrent une forte corrélation entre ces six descripteurs de texture et les paramètres de pénétrométrie. Les meilleures corrélations ($n=15$) pour les descripteurs *croquant*, *fondant*, *masticabilité* et *résistance au touché*, descripteurs qui caractérisent la fermeté sensorielle des pommes, sont obtenues pour F_f ($R=0,81$, $R=-0,81$, $R=0,74$ et $R=0,88$), W_f ($R=0,87$, $R=-0,85$, $R=0,83$ et $R=0,89$) et $Grad$ ($R=0,87$, $R=-0,81$, $R=0,78$ et $R=0,92$). Ces résultats sont en accord avec les conclusions de travaux antérieurs. Harker *et al.* (2002b) ont démontré l'efficacité des mesures de pénétrométrie pour prédire le *croquant* et la *fermeté* des pommes. Karlsen *et al.* (Karlsen *et al.*, 1999) ont également montré une forte corrélation entre les mesures de pénétrométrie et la *fermeté*. Dans une étude monovariétale de l'influence du génotype et de l'environnement, King *et al.* (2000) ont observé une bonne corrélation entre la fermeté mesurée par pénétrométrie et les descripteurs *fermeté* et *fondant*. La pénétrométrie est donc un bon prédicteur de la texture des pommes. Néanmoins, cette technique présente le désavantage d'être ponctuelle et donc de ne pas refléter la qualité du fruit entier. En effet, les mesures effectuées sur la face du fruit directement exposé au soleil conduisent à une moins bonne prédiction des descripteurs sensoriels (Harker *et al.*, 2002b). Ce phénomène n'est pas surprenant et peut être expliqué par les conséquences de la plus forte exposition au soleil. De plus, la destruction totale ou

partielle de l'échantillon présente un inconvénient majeur lorsque cette mesure est utilisée pour effectuer des mesures fruit à fruit en parallèle d'une autre mesure ou pour des mesures en ligne.

5.1.3.3 *Mesure instrumentale non destructive de la texture*

La sonométrie est une technique simple basée sur les propriétés vibrationnelles des produits. Les ondes sonores et ultrasonores interagissent avec la matière et peuvent être transmises, réfléchies, réfractées ou diffractées. La vitesse de propagation des ondes, l'atténuation et la réflexion sont des paramètres importants pour caractériser les propriétés des tissus des fruits. Lorsqu'un objet est excité à une fréquence, il entre en vibration. Certaines fréquences impliquent une vibration plus grande et provoquent des pics d'amplitude importante. L'état de la matière à ces fréquences particulières est appelée résonance. Les fréquences de résonance sont liées à l'élasticité, aux frictions internes, à l'humidité, à la taille, à la forme et à la densité de la matière. Pour des produits de même taille et de même forme, plus la chair est ferme, plus la fréquence de résonance est élevée. De plus, il est bien connu que la maturité de fruits comme le melon peut être vérifiée en écoutant le son produit lorsqu'on tape le fruit. La sonométrie suit ce principe. Dans une première étude sur la relation entre les propriétés vibrationnelles des pommes et leurs caractéristiques sensorielles, un indice de fermeté acoustique ($S = f^2 \times m^{2/3}$, avec f la première fréquence de résonance et m la masse du fruit) a été corrélé aux descripteurs sensoriels suivant : *texture globale*, *fermeté*, *résistance* et *jutosité* (Finney, 1971). L'indice de fermeté acoustique utilisé montrait une corrélation positive et significative avec les descripteurs de texture (e.g. pour la période d'octobre 1969 ($n=48$) : *texture globale* : $R=0,66$, *fermeté* : $R=0,68$, *résistance* : $R=0,61$, *jutosité* : $R=0,58$). Plus tard, d'autres études ont confirmé ces premières observations. Abbott *et al.* (1992) ont démontré un lien entre la fermeté sensorielle et la fermeté acoustique ($R=0,97$; $n=22$). Six sujets ont évalué la maturité de pommes Delicious. L'évaluation de la maturité était multicritère : observation, pression de la main, du pouce, dégustation. Néanmoins, la fermeté sensorielle était le critère primordial déclaré par les sujets et l'échelle de maturité se basait sur la fermeté des fruits. Barreiro *et al.* (1998) ont montré que la mesure de fermeté acoustique était fortement liée aux descripteurs sensoriels caractérisant la texture des fruits (*dureté*, *masticabilité*, *densité*, *granulosité* et *farinosité*). Par ailleurs, l'indice de fermeté acoustique a également été utilisé pour mesurer les changements de fermeté au cours du stockage (De Belie *et al.*, 2000) et de la maturation (Mehinagic *et al.*, 2006). Dans l'objectif de développement d'un outil de contrôle automatique, De Belie *et al.* (2000) avaient choisi d'étudier la mesure acoustique. Les auteurs ont montré la capacité de cet outil à détecter les changements de fermeté se produisant au cours du stockage de deux variétés – Cox et Elstar. De plus, une plus forte sensibilité du sonomètre, comparé au pénétromètre a été observée (De Belie *et*

al., 2000). Au cours de l'étude des relations entre mesures sensorielles et mesures instrumentales, Mehinagic *et al.* (2006) ont montré l'efficacité de la sonométrie à discriminer des niveaux de maturité – pré-maturité, maturité et post-maturité – pour trois variétés de pommes – Golden Delicious, Braeburn et Fuji. Par ailleurs, au regard des coefficients de corrélation observés entre les descripteurs sensoriels (*résistance au toucher, croquant*) et la mesure acoustique, ces auteurs ont émis des réserves quant à la possibilité de prédire les descripteurs sensoriels par la mesure acoustique.

5.2 La présentation des échantillons est un point clé en évaluation sensorielle

En évaluation sensorielle, la présentation des échantillons est un point important. En effet, les conditions de dégustation doivent être maîtrisées. Il est essentiel que la présentation des échantillons (taille, quantité) soit homogène pour tous les sujets. Le développement de méthodes d'évaluation sensorielle spécifiques aux produits végétaux a posé des interrogations quant au mode de présentation. Williams et Carter (1977), dans leur construction d'une méthode d'évaluation de la saveur des pommes Cox's Orange Pippin, se sont posés ces questions relatives à la présentation des échantillons : comment présenter l'échantillon : entier ou tranché? pelé ou non pelé ? Dans ces travaux, il est également souligné que la présentation en vrac au lieu d'un fruit par sujet pourrait assurer un échantillon plus uniforme. La littérature postérieure offre un grand nombre d'exemple de présentation des échantillons. Des exemples pour l'évaluation sensorielle des pommes sont présentés dans le Tableau 4 et sont détaillés dans les paragraphes suivants. Le mode de présentation des échantillons est un point clé qui dépend des objectifs de l'étude. La présentation peut être conditionnée par la présentation usuelle des produits (Abbott *et al.*, 2004). Deux autres raisons expliquant la présentation des échantillons ont été identifiées : les contraintes expérimentales et la volonté de réduire la variabilité entre échantillons.

Tableau 4. Présentation des échantillons de pomme en évaluation sensorielle

Présentation	Mode de distribution NC = Non communiqué	Raison de présentation NC = Non communiqué	Références
Cubes	NC	NC	(Aaby <i>et al.</i> , 2002)
Tranches	8 traitements testés, les tranches dégustées par un sujet proviennent de 2 fruits seulement	Limiter la variabilité fruit-à-fruit	(Abbott <i>et al.</i> , 2004)
	Trois tranches du même fruit pour les trois sujets	Limiter la variabilité fruit-à-fruit	(Brookfield <i>et al.</i> , 2011)
	Deux tranches du même fruit pour deux des quatre sujets	Limiter la variabilité fruit-à-fruit	(Oraguzie <i>et al.</i> , 2009)
1/8 pomme	Un fruit pour le groupe	Limiter la variabilité fruit-à-fruit	(Lateur <i>et al.</i> , 2001)
	3 1/8 fruit par sujet	NC	(Watada <i>et al.</i> , 1980)
1/4 pomme	Un fruit par sujet	NC	(Mehinagic <i>et al.</i> , 2003)
	4 traitements testés, les quartiers de fruit dégustés par un sujet proviennent d'un fruit unique	Limiter la variabilité fruit-à-fruit	(Bhagwat <i>et al.</i> , 2004)
	Un fruit par sujet	NC	(Kühn <i>et Thybo</i> , 2001)
1/2 pomme	Un fruit par sujet	Les deux moitiés étaient mesurées par le même sujet	(Mehinagic <i>et al.</i> , 2004)
	Division du fruit pour séparer les faces ensoleillée et non ensoleillée		
	Un fruit par sujet	Une moitié pour l'évaluation sensorielle et l'autre pour la mesure instrumentale	(Harker <i>et al.</i> , 2002b)
	Division du fruit pour obtenir deux moitiés similaires		
	Un fruit par sujet	NC	(King, G. J. <i>et al.</i> , 2000)

5.2.1 La présentation des échantillons répond souvent à des contraintes expérimentales

Harker *et al.* (2002b, 2002c) ont comparé les résultats de mesures instrumentales et sensorielles de la texture des pommes. Afin d'effectuer une comparaison fruit à fruit, chaque pomme était coupée en deux moitiés. L'une des moitiés de chaque fruit était évaluée par un sujet et la moitié correspondante était analysée instrumentalement. Les mesures instrumentales et la mesure sensorielle étaient donc réalisées sur le même fruit.

Mehinagic *et al.* (2004), dans le souci de prendre en compte les différences pouvant exister entre les faces ensoleillées et non ensoleillées des pommes, ont dupliqué chaque mesure, l'une des duplications était réalisée sur la face ensoleillée et l'autre sur la face non ensoleillée. Les mesures étaient ensuite moyennées. Il aurait pu être intéressant d'explorer l'effet face en moyennant les mesures par face.

5.2.2 La présentation des échantillons est parfois utilisées pour réduire la variabilité

Abbott *et al.* (2004), ont étudié l'acceptabilité de tranches de pommes fraîches. Quatre variétés ont été évaluées ainsi que deux traitements empêchant le brunissement et permettant le maintien de la fermeté (un nouveau traitement, PQSL et un traitement commercial, COM). Dans une étude préliminaire pour vérifier l'acceptation des concentrations des ingrédients du nouveau traitement, l'évaluation sensorielle de tranches de pommes traitées a été réalisée par dix sujets. Chaque tranche traitée était comparée à une tranche non traitée provenant de la même pomme, limitant ainsi la variabilité due aux fruits. Ainsi, pour chaque sujet les variations fruit à fruit étaient limitées. Cette même stratégie a été utilisée par Bhagwat *et al.* (2004). Ce protocole permet de réduire la probabilité d'attribuer la différence perçue entre une tranche traitée et une tranche non traitée au traitement alors qu'elle est due à des différences entre les deux pommes dont sont issues les tranches.

Lateur *et al.* (2001) ont, quant à eux, limité les variations fruit à fruit entre les sujets. Leur panel était composé de huit sujets et pour chaque échantillon dégusté, les huit sujets recevaient un huitième d'une même pomme. De même, Brookfield *et al.* (2011) et Oraguzie *et al.* (2009) ont utilisé le partage des fruits pour homogénéiser les échantillons proposés aux sujets. Dans ces deux études, l'homogénéisation avait pour but d'éliminer les variations dues aux fruits afin de pouvoir comparer les évaluations de plusieurs sujets.

Aucun des auteurs cités précédemment n'a étudié l'impact de la stratégie de présentation des échantillons sur la qualité des mesures sensorielles obtenues ou tenté de quantifier la part des différentes sources de variabilité. Partager un fruit pour limiter la variabilité biologique paraît évident. Néanmoins, nous avons vu précédemment que la variabilité biologique s'exprime également au sein des pommes, entre les faces opposées et entre les parties supérieure et inférieure du fruit. À notre connaissance, seule une étude comparant les mesures instrumentales et sensorielles de la qualité des fraises explore l'impact du mode de présentation sur la qualité des résultats sensoriels (Gunness *et al.*, 2009). Deux présentations des fraises pour l'évaluation sensorielle ont été étudiées : fruit individuel et purée de fruit. Les compositions de la variation totale des notes sont comparées. Les variations entre les fruits individuels sont généralement plus grandes que les variations entre les échantillons individuels de purée. Pour la saveur sucrée, la variation de chacun des fruits contribue pour 63 % à la variation totale, tandis que la contribution des échantillons sous forme de purée n'est que de 31 %. L'évaluation de purée permet donc une réduction de la variabilité due aux produits. De plus, les différences entre niveaux de maturité sont mieux affirmées dans le cas d'échantillons de purée. Cependant, la présentation sous forme de purée ne permet pas l'évaluation de la texture et peut impliquer des modifications de la flaveur. En outre, la diminution de la variance pour la présentation sous forme de purée pourrait être due à une atténuation des différences par le processus de brassage ou à une réduction de la complexité de l'évaluation étant donné que les interactions avec la texture n'existent plus.

5.3 Bilan sur la maîtrise de la variabilité

La connaissance de la variabilité des caractéristiques des pommes a conduit à l'utilisation de méthodes d'échantillonnage et de présentation des fruits pour tenter de réduire la variabilité des lots étudiés. L'efficacité de ces méthodes n'a pas ou peu été démontrée. Or, certaines méthodes peuvent être chronophage et/ou induire des contraintes d'évaluation. En effet, le temps nécessaire à la mesure instrumentale des fruits préalablement à l'échantillonnage peut être important ; et, les dégustations de purées, cubes, tranches ou fruits entiers ne permettent pas des évaluations équivalentes en termes de qualité des descripteurs et du nombre de descripteurs.

6. Conclusion

La réponse sensorielle est le résultat d'un processus complexe et comporte une forte variabilité. Les variations dues aux sujets d'un panel sont l'une des causes de cette variabilité. En effet, des

variations inhérentes génétiques et liées à l'interaction entre le sujet et son environnement existent. La standardisation des méthodes de mesure, et notamment du profil conventionnel dans le cadre de la caractérisation de produits, tend à réduire les variations indésirables et fait donc de cette mesure une mesure objective.

Les variations inhérentes aux sujets ne sont cependant pas totalement éliminées. Malgré un entraînement dédié, les sujets varient dans leurs seuils de perception et dans leur utilisation de l'échelle de notation (moyenne de notation, dispersion des notes, désaccord). Le développement de méthodes spécifiques pour l'analyse des données sensorielles a donc été entrepris afin d'intégrer ces variations connues. Les méthodes développées sont cependant peu utilisées et le modèle d'analyse de la variance à deux facteurs produit et sujet et leur interaction est utilisé de façon standard.

De plus, lors de l'évaluation sensorielle de pommes, les variations entre les fruits d'un lot ont été mises en avant comme l'une des causes possibles des variations observées entre les sujets. Ces variations sont en effet importantes puisqu'un lot est constitué de différents fruits provenant de différents arbres et probablement de différents vergers et que des différences ont été observées à toutes ces échelles. En outre, les différences perceptibles au sein d'un fruit viennent grossir la variabilité observée des résultats sensoriels.

L'observation des variations d'une réponse constituent l'objectif de l'analyse de la variance qui est une méthode très utilisée pour l'analyse des données sensorielles. Les modèles sont malléables et peuvent être adaptés à la structure des données. Dans le cadre de l'évaluation sensorielle des pommes, une adaptation du modèle d'analyse de la variance pour une prise en compte de la variabilité intra-lot serait judicieuse.

En évaluation sensorielle, diverses stratégies ont été utilisées pour réduire la variabilité des lots dégustés, souvent perçue comme une nuisance. L'homogénéisation par mesure instrumentale ou par une présentation des échantillons spécifique a été employée. Cependant, l'impact de ces stratégies sur la variabilité de la réponse sensorielle n'a été observé que pour une étude. L'efficacité des stratégies avancées reste alors à prouver.

Partie II. Problématique et démarche expérimentale

1. Problématique et hypothèses

L'étude bibliographique a montré que la variabilité des données sensorielles pommes est au centre des préoccupations des chercheurs et professionnels utilisant l'analyse sensorielle pour évaluer la qualité des fruits (e.g. Hampson *et al.* (2000) et Seppä *et al.* (2012)). Cette variabilité est due d'une part aux variations entre les sujets du panel, et d'autre part aux variations entre les fruits d'un lot. Le facteur sujet a été l'objet de nombreuses études et cette question est encore traitée aujourd'hui (Brockhoff, 2003, Brookfield *et al.*, 2011, Oraguzie *et al.*, 2009, Romano *et al.*, 2008). Concernant le facteur fruit, il a été très souvent présenté comme un challenge et de nombreuses études ont démontré la présence de variabilité biologique à différentes échelles.

La variabilité biologique des pommes est devenue un sujet majeur d'étude pour les spécialistes de la qualité des fruits (Bobelyn *et al.*, 2010, De Ketelaere *et al.*, 2006, Hertog, 2002, Hertog *et al.*, 2004, Schouten *et al.*, 2004, Tijssens *et al.*, 2003) mais n'a pas été étudiée dans le cadre de la qualité sensorielle. Or, la mesure et le contrôle de la variabilité biologique nous semblent essentiels pour obtenir des résultats sensoriels pertinents et de qualité. La problématique de cette thèse repose donc sur la recherche de méthodologies permettant de mesurer et gérer la variabilité intra-lot et la variabilité des sujets exprimées dans une réponse sensorielle complexe et de les prendre en compte dans les traitements afin d'améliorer la qualité de la mesure sensorielle.

Avant de nous engager dans l'amélioration des méthodologies d'évaluation sensorielle, nous avons jugé indispensable de vérifier la sensibilité du panel et plus précisément sa capacité à discriminer des fruits d'un même lot d'une variété donnée présentant des différences biologiques. Nous avons procédé à une mesure de performance en deux temps pour vérifier les hypothèses de départ suivantes :

- Un panel entraîné à l'évaluation sensorielle des pommes est capable de discriminer les fruits de variétés distinctes (Figure 6, étape I-1) ;
- Un panel entraîné à l'évaluation sensorielle des pommes est aussi capable de discriminer les fruits d'une même variété (un lot) présentant des différences biologiques (Figure 6, étape I-2).

La validation des capacités de discrimination du panel au niveau inter-lot (inter-variété ici) mais aussi intra-lot a mené à considérer deux approches permettant d'améliorer la qualité de la réponse sensorielle.

La première approche portait sur l'analyse des données issues de profils sensoriels de produits variables. L'objectif était d'adapter le traitement des données par la prise en compte de toutes les sources de variation intervenant dans la réponse sensorielle pomme, et notamment la variabilité intra-lot. Les hypothèses suivantes ont été vérifiées :

- Dans les données sensorielles, la contribution des variations intra-lot prend une part importante (Figure 6, étape II-1) ;
- Il existe des différences d'utilisation de l'échelle entre les sujets qui sont de trois types : notation moyenne, dispersion et désaccord (Figure 6, étape II-2-1);
- L'intégration d'un facteur relatif à la variabilité intra-lot et l'application d'une approche de type « assessor » dans les modèles modifient sensiblement les résultats de discrimination entre les lots (Figure 6, étapes II-1 et II-2-2).

La deuxième approche portait sur l'impact de la réduction de la variabilité des fruits dégustés par le panel sur la variabilité de la réponse sensorielle. Le postulat de ce travail est que la variabilité biologique existant au sein d'un fruit est inférieure à la variabilité biologique entre les fruits d'un même lot. L'hypothèse à vérifier est la suivante :

- En diminuant la variabilité intra-lot des lots dégustés, par une méthode d'échantillonnage ou de présentation des échantillons, il est possible de diminuer la variabilité des données sensorielles (Figure 6, étape III-1 et III-2).

2. Démarche expérimentale

Un panel de 19 sujets a été recruté et formé à l'évaluation sensorielle des pommes. Trois variétés de pommes ont été étudiées : Ariane, Braeburn et Pink Lady®. Les trois variétés choisies ont des caractéristiques proches (Vaysse *et al.*, 2000). Ce choix a été guidé par la volonté de pouvoir observer la capacité de discrimination du panel entre des variétés (lots) mais aussi de pouvoir observer l'impact des variations intra-lot sur cette capacité de discrimination intra-lot. En effet, l'impact d'une forte variabilité intra-lot sur les résultats de lots très différents (e.g. pour le descripteur *croquant*, la Reine des Reinettes, peu croquante, et la Granny Smith, très croquante) est supposé moins important que pour des lots plus proches (e.g. pour le descripteur *croquant*, la Pink Lady® et la Granny Smith, très croquantes).

2.1 *Vérification de la sensibilité du panel*

Pour commencer, après une phase de sélection et d'entraînement, nous avons vérifié la capacité du panel à discriminer ces trois variétés de pommes (Figure 6, étape I-1). Puis, nous avons testé la capacité du panel à faire des différences fines au sein de chaque variété entre des sous-lots présentant des différences biologiques et échantillonnés par une méthode non destructive. (Figure 6, étape I-2). Ces résultats font l'objet du Chapitre I de la Partie IV.

2.2 *Adaptation du traitement des données sensorielles pommes*

Une fois la capacité des sujets à faire des différences sensibles au sein des lots validée, nous souhaitons identifier les facteurs responsables de la variabilité de la réponse sensorielle et quantifier les contributions de chacun de ces facteurs. L'analyse de la variance est couramment utilisée pour le traitement des données sensorielles car elle fournit des informations détaillées sur chaque descripteur et sur la significativité des effets des facteurs étudiés. Les modèles mixtes permettent une caractérisation simultanée des effets principaux qui influencent l'expérience, ainsi que l'estimation des différentes composantes qui influent sur la variabilité intrinsèque totale (Pinheiro et Bates, 2000). Par conséquent, le travail visant à mesurer la variabilité s'est focalisé sur l'utilisation de ces modèles mixtes dans l'analyse de la variance.

Dans un premier temps, nous avons étudié un modèle permettant de tenir compte de la variabilité biologique des fruits, adapté à partir du modèle standard d'analyse de la variance (Figure 6, étape II-1). L'efficacité de ce modèle a été évaluée en comparant les résultats à ceux du modèle standard. Ces

résultats font l'objet du Chapitre II.1 de la Partie IV. Le modèle ainsi déterminé est adapté à l'analyse de données sensorielles issues de profils de produits présentant une forte variabilité.

Puis, suite à la caractérisation des tendances de notation de chaque sujet pour chaque descripteur et chaque variété et l'évaluation de l'accord entre les sujets (Figure 6, étape II-2-1), une nouvelle approche du modèle tenant compte de la variabilité intra-lot, basée sur le modèle « assessor », a été étudiée. Dans cette approche, le terme d'interaction a été décomposé en désaccord pur et en dispersion (Figure 6, étape II-2-2). Le Chapitre II.2 de la Partie IV est consacré à l'approche « assessor » du modèle.

2.3 Réduction de la variabilité des fruits dégustés

Parallèlement à l'approche précédente, nous souhaitons identifier des méthodes permettant *a priori* de réduire la variabilité de la réponse sensorielle en diminuant celle due aux fruits. En effet, la variabilité des sujets n'est pas seule responsable de la variabilité globale de la réponse sensorielle. Pour ce faire, deux stratégies ont été identifiées.

D'une part, d'après la littérature, certaines mesures instrumentales sont de bons outils de prédiction des caractéristiques sensorielles. Deux mesures instrumentales ont été choisies pour l'homogénéisation (Figure 6, étape III-1), la première étant une mesure de référence mais endommageant le fruit : la pénétrométrie ; et la seconde étant non destructive mais moins bien corrélée aux caractéristiques sensorielles : la sonométrie. Pour chaque mesure instrumentale et chaque variété, deux lots, l'un hétérogène et l'autre homogène, ont été prélevés du lot initial puis évalués par le panel. Pour chaque mesure instrumentale testée, les écart-types des données sensorielles ont été calculés pour chaque variété et chaque échantillonnage, puis comparés. Les différences entre les résultats d'analyse de la variance pour le profil sur lots homogènes et le profil sur lot hétérogènes ont été étudiées. Ces résultats font l'objet du Chapitre III.1 de la Partie IV.

D'autre part, il semble intuitif que le partage des fruits entre plusieurs sujets permettrait de réduire la part de la variabilité globale issue de la variabilité biologique. Certains auteurs ont appliqué cette méthode sans en vérifier l'efficacité. Nous avons donc choisi trois méthodes de présentation des fruits : un fruit pour un sujet, un fruit pour quatre sujets et un fruit pour huit sujets, avec à chaque fois l'évaluation d'un huitième de fruit (Figure 6, étape III-2). Les écart-types des données ont été calculés pour chaque variété et chaque méthode de présentation, puis comparés. L'impact du partage des fruits entre les sujets sur les résultats d'analyse de la variance a été étudié. Ces résultats font l'objet du Chapitre III.2 de la Partie IV.

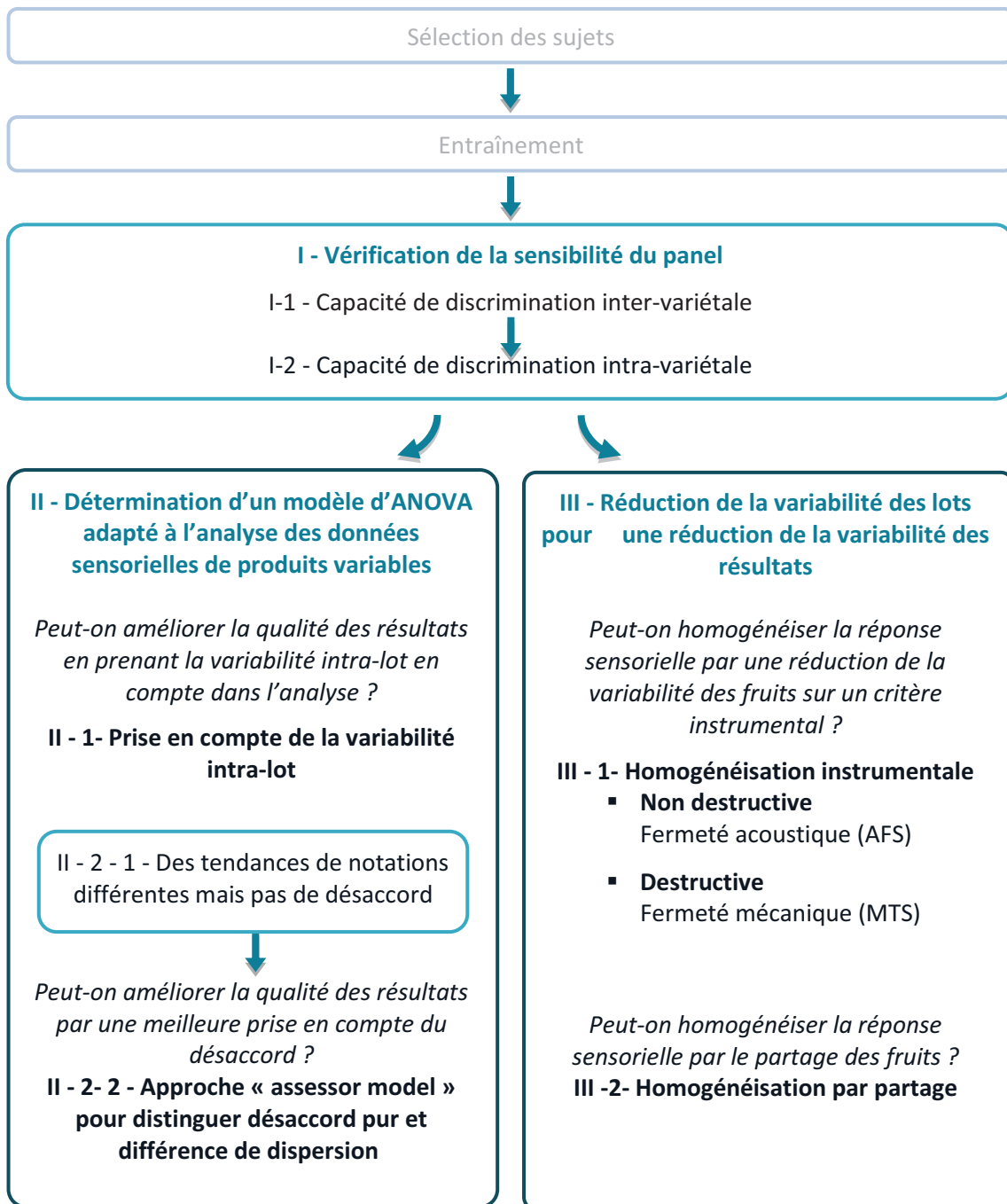


Figure 6. Démarche expérimentale

Partie III. Matériels et méthodes

1. Notre produit modèle : la pomme

Trois variétés de pommes ont été choisies : Ariane, Braeburn et Pink Lady®. Les fruits ont été achetés à un grossiste angevin, Les Vergers d'Anjou. Chaque année, les fruits provenaient d'un producteur ayant récolté sur une journée en fonction de la date optimale de maturité commerciale pour chaque variété étudiée. Les fruits ont été stockés en chambre à 4°C sous atmosphère normale (AN) ou contrôlée (AC). Les caractéristiques des pommes sont présentées dans le Tableau 5. Pour répondre à nos besoins expérimentaux, une partie des pommes dégustées de novembre à décembre 2010 a été conservée à une température supérieure ($18^{\circ}\text{C} \pm 2^{\circ}\text{C}$).

Tableau 5: Caractéristiques des pommes utilisées

Variété	Année de récolte	Calibre (mm)	Stockage	Période de dégustation
Ariane	2010	70/75	AN 4°C et 18°C	Novembre à décembre 2010
Braeburn	2010	70/75	AN 4°C et 18°C	Novembre à décembre 2010
Pink Lady®	2010	70/75	AN 4°C et 18°C	Novembre à décembre 2010
Ariane	2010	70/75	AC 4°C puis AN 4°C	Janvier à juin 2011
Braeburn	2010	70/75	AC 4°C puis AN 4°C	Janvier à juin 2011
Pink Lady®	2010	70/75	AC 4°C puis AN 4°C	Janvier à juin 2011
Ariane	2011	70/75	AN 4°C	Novembre 2011 à février 2012
Braeburn	2011	70/75	AN 4°C	Novembre 2011 à février 2012
Pink Lady®	2011	70/75	AN 4°C	Novembre 2011 à février 2012

2. La caractérisation sensorielle

2.1 *Le panel entraîné*

Le panel sensoriel comprenait 19 sujets – 16 femmes et 3 hommes, âgés de 33 à 63 ans – recrutés en avril et mai 2010 pour les besoins de l'étude. Ces sujets ont été formés à l'évaluation de la texture et des saveurs des pommes jusqu'en Octobre 2010, à raison de deux séances de 1,5 heures par semaine (soit 40 heures environ d'entraînement). La sélection et la formation des sujets ont été effectuées conformément aux recommandations des normes ISO (ISO, 1993). Les descripteurs sensoriels étudiés étaient *croquant, fermeté, crissant, jutosité, fondant, acidité et sucré* tels que définis dans le Tableau 6. Ces descripteurs ont été générés et définis par l'ensemble des sujets et correspondent aux descripteurs généralement utilisés pour la description des pommes (Harker *et al.*, 2002b, Harker *et al.*, 2002c, Mehinagic *et al.*, 2004). Une fois les descripteurs sensoriels bien définis et compris, les sujets ont été formés à la détection de petites différences à travers des exercices spécifiques. La comparaison par paires (ISO, 2005a) et le profil sensoriel avec retour immédiat sur performance (Findlay *et al.*, 2007) ont notamment été utilisés. Les sept descripteurs étaient évalués dans l'ordre indiqué dans le Tableau 6, sur une échelle continue allant de la plus faible intensité (valeur 0) à la plus forte intensité (valeur 10). Les descripteurs étaient évalués sur huitième(s) de pomme, présenté(s) non pelé(s). Étant donné l'influence de l'orientation de la morsure lors de la dégustation de pommes (Khan et Vincent, 1993), les instructions de dégustation étaient précises. Il était demandé aux sujets de tenir le huitième de pomme entre le pouce et l'index avec la peau du fruit vers l'extérieur, puis de croquer le morceau de pomme sans atteindre la peau. Les sujets avaient pour instruction de se rincer la bouche entre chaque échantillon. Pour cela, de l'eau et des crackers non salés étaient mis à leur disposition. Pour chaque évaluation, les sujets recevaient un ou deux huitième(s) de pomme. Dans le cas où seulement un huitième de pomme était présenté, la liste de descripteurs était réduite à cinq descripteurs – *croquant, fermeté, jutosité, fondant et acidité*.

2.2 *Les conditions de dégustation sont standardisées*

Les profils sensoriels ont toujours été effectués dans des conditions contrôlées. Les sujets évaluaient les fruits dans des cabines équipées de lumières rouges pour limiter l'influence de la couleur de la peau, dans une salle répondant aux normes ISO (ISO, 2007a) dont la température était fixée à 21°C

$\pm 1^\circ\text{C}$. Les échantillons étaient présentés de manière monadique séquentielle. L'ensemble des résultats a été acquis à l'aide du logiciel Fizz réseau (version 2.46 ; Biosystems, Courtenon, France).

Tableau 6. Descripteurs sensoriels utilisés par le panel

Descripteurs	Définition et protocole d'évaluation	Références
Croquant	Bruit provoqué par l'enfoncement des incisives dans la chair et par le détachement du morceau de pomme lorsqu'on le mord <i>Enfoncez franchement les incisives dans le quartier de pomme pour détacher un morceau et évaluez le bruit émis</i> <i>Prise en bouche du morceau de pomme</i>	Borne - : Golden Delicious stockée à T_{amb} pendant 2 semaines Borne + : Pink Lady® stockée à 4°C
Fermeté	Résistance de la chair à la pression des molaires lors de la première mastication <i>Écrasez le morceau de pomme entre les molaires et évaluez la force nécessaire pour enfoncer les molaires</i> <i>1^{ère} mastication</i>	Borne - : Elstar stockée à 4°C Borne + : Granny Smith stockée à 4°C
Crissant	Bruit provoqué par le broyage de la pomme entre les dents <i>Évaluez le bruit émis par le broyage de la pomme par les dents pendant 3 à 4 mastications du morceau de pomme</i> <i>Pendant 3-4 mastications</i>	Borne - : Golden Delicious stockée à T_{amb} pendant 2 semaines Borne + : Tentation stockée à 4°C
Jutosité	Quantité de jus libéré en bouche <i>Mastiquez le morceau de pomme 4 à 5 fois, sans avaler, puis évaluez la quantité de jus en bouche en avalant le jus</i> <i>Après 4-5 mastications</i>	Borne - : Golden Delicious stockée à 4°C Borne + : Pink Lady® stockée à 4°C
Fondant	Capacité des morceaux de pommes à se dissoudre plus ou moins rapidement dans la bouche sous l'effet de la salive <i>Évaluez la quantité de morceaux de pommes et leur taille après 5 à 6 mastications</i> <i>Après 5-6 mastications</i>	Borne - : Golden Delicious stockée à 4°C Borne + : Golden Delicious stockée à T_{amb} pendant 2 semaines
Acidité	Intensité de la saveur acide perçue en bouche lors de la mastication <i>Évaluez l'intensité de la saveur acide pendant toute la dégustation</i> <i>Pendant les mastications 1-6</i>	Borne - : 1.3g/L acide malique et 68g/L saccharose Borne + : 8.0 g/L acide malique et 68g/L saccharose
Sucré	Intensité de la saveur sucrée perçue en bouche lors de la mastication <i>Évaluez l'intensité de la saveur sucrée pendant toute la dégustation</i> <i>Pendant les mastications 1-6</i>	Borne - : 68g/L saccharose et 1.3g/L acide malique Borne + : 140g/L saccharose et 1.3g/L acide malique

3. La mesure instrumentale de la texture

Les mesures instrumentales ont été utilisées pour la sélection des fruits. Les mesures étaient donc faites sur fruit entier avant l'évaluation sensorielle. Comme indiqué précédemment, la pénétrométrie et la sonométrie ont été utilisées.

3.1 *La fermeté mécanique, une mesure de référence*

La pénétrométrie est une technique simple et rapide. C'est la méthode de référence pour la mesure de la fermeté des fruits (Harker *et al.*, 2002a). Cette mesure est basée sur l'introduction d'un module cylindrique en inox dans la chair d'un fruit pelé pour mesurer la force de pénétration maximale. Cette mesure peut être réalisée avec des appareils manuels ou automatisés, ces derniers étant plus reproductibles. Duprat *et al.* (Duprat *et al.*, 2000) ont proposé un protocole avec un pénétromètre automatisé, consistant à enregistrer la force nécessaire au déplacement d'une sonde dans les dix premiers millimètres du fruit. Leur protocole permet de mesurer la résistance de l'épiderme et de la chair. La courbe force/déformation obtenue pour un fruit non pelé (Figure 7) comporte une phase de compression linéaire présentant un maximum, une dépression brutale et une phase augmentant légèrement correspondant aux propriétés de la chair (Duprat *et al.*, 2000). Parmi, les sept paramètres identifiés sur la courbe force/déformation (Duprat *et al.*, 2000) (Figure 7), la force mesurée à 7 mm de déplacement (F_f), comparable à la fermeté Magness-Taylor du fruit épluché, couramment utilisée par les professionnels de la filière arboricole, a été sélectionnée pour servir ces travaux de thèse.

Les mesures de pénétrométrie ont été effectuées sur pommes entières non épluchées avant la dégustation, avec une machine de traction universelle (MTS Universal, Synergie 200H) selon la méthode décrite par Duprat *et al.* (2000). L'appareil fonctionne à l'aide d'un moteur pas à pas permettant une vitesse de progression de la sonde constante. Une sonde cylindrique métallique à extrémité hémisphérique de 4 mm de diamètre a été utilisée pour la pénétration du fruit. La pénétration de la sonde était fixée à 10 mm. Le fruit était perforé au niveau équatorial à une vitesse constante de 50 mm.min⁻¹. Chaque fruit a été mesuré en un point afin de ne pas trop détériorer le fruit qui était destiné à être dégusté.

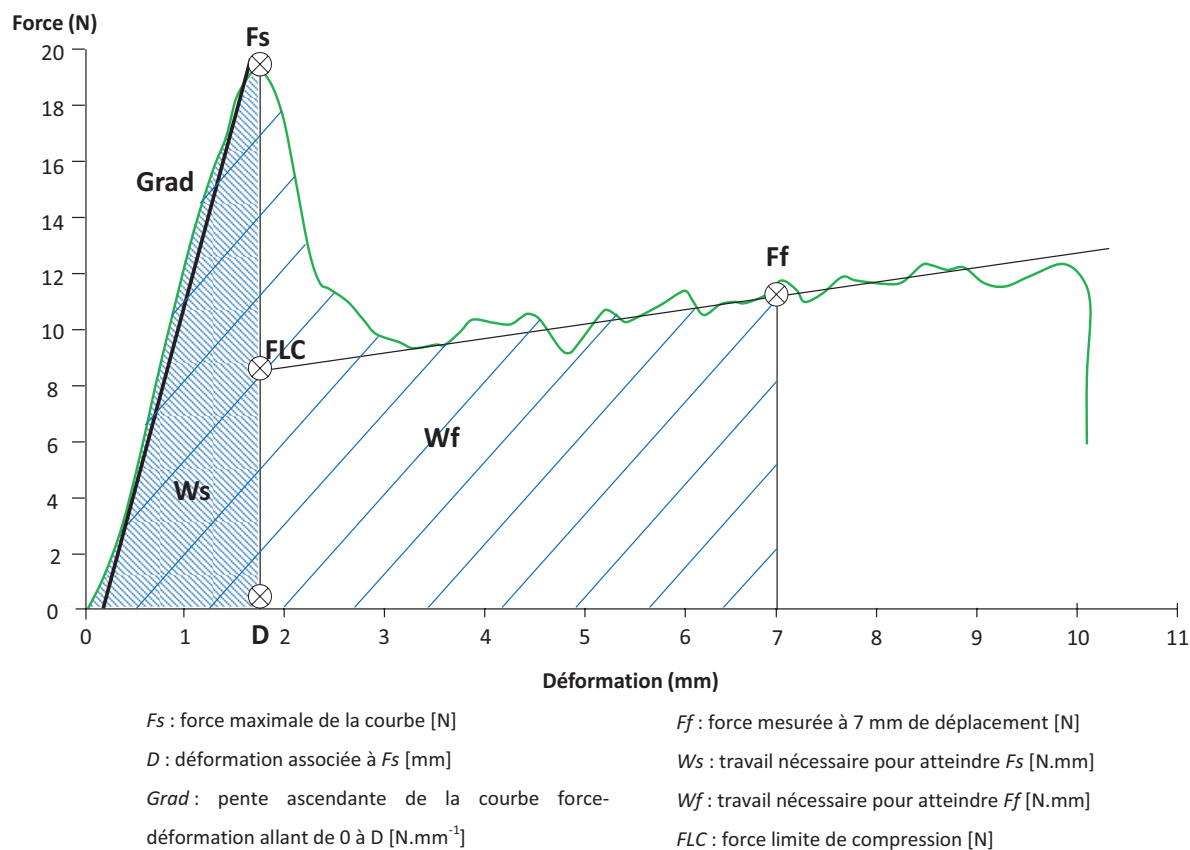


Figure 7. Courbe force/déformation obtenue par pénétrométrie sur pomme non épluchée en utilisant la machine de traction universelle (Mehinagic, 2004)

3.2 La fermeté acoustique, une mesure non-destructive

La sonométrie est une technique basée sur les propriétés vibrationnelles des produits. Les mesures ont été effectuées avec l'appareil AWETA Acoustic Firmness Sensor (AFS; AWETA, Nootdorp, The Netherlands) (Figure 8). Il s'agit d'un périphérique commandé par un ordinateur. Le dispositif de mesure est constitué d'un boîtier contenant un système de pesée, une bille et un microphone. Le boîtier est équipé d'un support en mousse pour recevoir le fruit à mesurer (Figure 8). Lorsque le fruit est déposé sur le support, sa masse est évaluée. Puis, un signal sonore est généré par un impact modéré de la bille sur le fruit. Le son émis est enregistré par un microphone. La bille est très légère de sorte que l'énergie dissipée dans le fruit, lors de l'impact, soit très faible et n'altère pas le fruit. Tous les signaux véhiculés sont analogiques. La réponse sonométrique se traduit par la courbe de réponse temporelle exprimant l'amplitude de vibration sonore en fonction du temps. À partir de cette réponse brute et par la transformée de Fourier, une deuxième courbe est obtenue : la réponse fréquentielle (amplitude = f [fréquence]). Elle exprime la contribution de chaque fréquence à la réponse temporelle par l'amplitude qui lui est associée. La fréquence de résonance (f_0) correspond

à la fréquence pour laquelle l'amplitude de vibration sonométrique est maximale. Elle est utilisée pour calculer l'indice de fermeté acoustique (S , *stiffness* en anglais) (Lu et Abbott, 1997, Shmulevich *et al.*, 2003), donné par l'équation (III-1).

$$S = f_0^2 * m^{2/3} \quad (III-1)$$

Avec f_0 = la fréquence de résonance [Hz] et m = la masse du fruit [g].

L'équation (III-1) est valable pour les objets sphériques (Zude *et al.*, 2006).

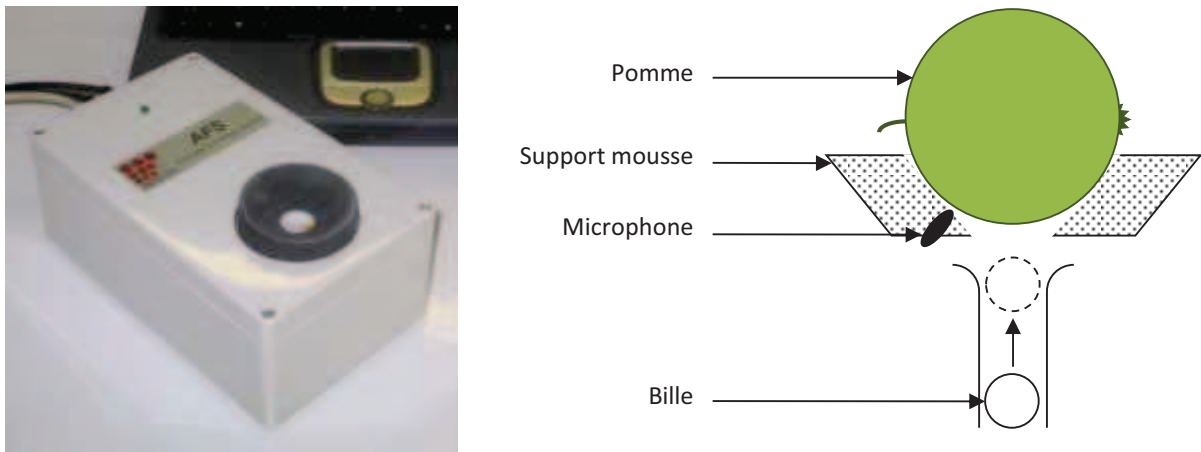


Figure 8. Sonomètre (AFS; AWETA, Nootdorp, The Netherlands) (à gauche) et schéma du dispositif (à droite)

Les mesures acoustiques ont été effectuées sur pommes entières non épluchées. Suivant les recommandations de Chen et DeBaerdemaeker (1993), les mesures ont été répétées quatre fois sur le pourtour du fruit avec un pas de 90° (Figure 9) puis moyennées.

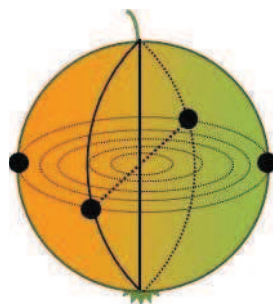


Figure 9. Emplacements des mesures acoustiques

4. Traitements statistiques

4.1 *L'analyse de la variance*

L'analyse de la variance (ANOVA) est la méthode statistique la plus utilisée en évaluation sensorielle. Plusieurs modèles ont été utilisés au cours du travail de thèse et comprennent plusieurs facteurs et leur(s) interaction(s). Les modèles utilisés sont des modèles mixtes comportant des effets fixes et des effets aléatoires. Lorsque les données sont équilibrées et présentent une structure produit à une dimension, l'analyse du modèle mixte peut être traitée par le simple calcul des carrés moyens combiné à l'utilisation du terme d'erreur approprié provenant d'une décomposition d'ANOVA. Cependant, les données générées au cours du travail de thèse ne sont pas équilibrées et présentent une structure en bloc complexe et/ou une structure produit à deux dimensions. Dans ce cas, la méthodologie décrite précédemment n'est pas adaptée. En effet, l'estimation de l'effet produit ne peut plus être réduite à la simple moyenne des données. Quant au choix du terme d'erreur, il devient compliqué car celui-ci devient une combinaison complexe des différentes composantes aléatoires. Les tests de significativité, les tests post-hoc (comparaison de moyenne) et l'estimation des composantes de la variance sont décrits ci-après.

4.1.1 Test de significativité

Le test F de Fisher a été utilisé pour tester la significativité des effets fixes (Pinheiro et Bates, 2000). Le calcul des degrés de liberté du dénominateur était effectué selon l'approximation de Satterthwaite. L'approximation de Satterthwaite permet de construire un pseudo test F de Fisher dans les modèles d'analyse de la variance pour lesquels un test F exact ne peut s'obtenir comme rapport de deux carrés moyens. Dans ces cas, il est possible d'utiliser une combinaison linéaire de carrés moyens au numérateur, au dénominateur ou aux deux endroits de telle sorte que les espérances du numérateur et du dénominateur soient égales sous l'hypothèse nulle.

Concernant les effets aléatoires, ils ont été testés en utilisant le maximum de vraisemblance restreint (en anglais, REML pour REstricted Maximum Likelihood). La vraisemblance est une mesure de la qualité d'ajustement des paramètres d'un modèle aux données. Il est naturel de calculer les estimateurs des paramètres du modèle de manière à maximiser la vraisemblance (maximum de vraisemblance). Le maximum de vraisemblance restreint est mieux adapté aux modèles mixtes car il permet de séparer les deux types de paramètres (fixe et aléatoire). Pour comparer deux modèles imbriqués, il est pertinent de comparer leur qualité d'ajustement. Les modèles A et B sont dit

imbriqués si le modèle B comporte un effet supplémentaire x par rapport au modèle A. L'hypothèse H_0 : le facteur x n'a pas d'influence sur la réponse, est testée en comparant les vraisemblances V_A et V_B des deux modèles. Sous l'hypothèse nulle, la statistique de test $-2\ln(V_A/V_B)$, aussi appelée *déviante*, suit une loi de Chi^2 à un degré de liberté (Pinheiro et Bates, 2000). Dans le cas d'un modèle comportant n effets aléatoires, nous procédons pas à pas. Chaque effet aléatoire du modèle le plus complet est testé individuellement. L'effet le moins significatif est éliminé et la procédure est répétée avec les $n-1$ effets aléatoires restants. Le processus est répété jusqu'à ce qu'il n'y ait plus d'effet aléatoire.

En pratique, les modèles ont été ajustés avec la fonction `lmer` du package `lme4`. Les effets ont ensuite été testés grâce à la fonction `step` du package `lmerTest` (Kuznetsova *et al.*, 2012) disponible en accès libre sous le logiciel R (version 2.15.1) (R Core Team, 2012). Les arguments de cette fonction sont le modèle linéaire à analyser, les données associées, la spécification des tests post-hoc, un niveau de significativité pour les effets fixes ($\alpha=0,05$ par défaut), un niveau de significativité pour les effets aléatoires ($\alpha=0,10$ par défaut) ainsi que deux arguments précisant si le modèle doit être réduit. Les sorties de la fonction `step` sont les p-values des effets aléatoires, les p-values des effets fixes, les moyennes estimées des niveaux des effets fixes et leurs différences, ainsi que le modèle final réduit.

Supposons un tableau de données nommé 'donnee', résultats de l'évaluation de I produits par J sujets en R répétitions pour une liste de K descripteurs. L'analyse du modèle standard d'ANOVA à deux facteurs produit (fixe) et sujet (aléatoire) et leur interaction (aléatoire), pour le descripteur k , est effectuée selon le code suivant :

```
> res <- lmer(donnee[,k] ~ Produit + (1|Sujet) + (1|Sujet:Produit),
             data=donnee)
> tres <- step(model=res, data=donnee, test.offs=c("Produit"), alpha.rand =
             0.10, alpha.fix = 0.05, isFixReduce = FALSE, isRandReduce =
             FALSE)
```

4.1.2 Comparaison de moyenne

La fonction `step` fournit les résultats de tests de comparaison de moyenne pour les effets spécifiés par l'argument `test.offs`. Le résultat de ces tests est enregistré dans le tableau `Differences`

o f LSMEANS. L'intervalle de confiance pour l'estimateur de la moyenne μ d'un échantillon de taille n et d'écart-type σ est défini comme suit :

$$\left[\bar{x} - t_{\alpha/2} \frac{\sigma}{\sqrt{n}}; \bar{x} + t_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \quad (\text{III-2})$$

où \bar{x} est la moyenne observée et $t_{\alpha/2}$ est le quantile d'ordre $\alpha/2$ de la loi de Student à $n - 1$ degrés de liberté.

4.1.3 Estimation des composantes de la variance

La décomposition de la variabilité est au cœur de nos préoccupations. Afin d'observer la répartition des différentes sources de variation testées, nous appliquons un modèle à effets aléatoires, c'est-à-dire que quel que soit le modèle mixte analysé, ce même modèle est appliqué de nouveau avec tous les effets aléatoires (King, G. J. *et al.*, 2000). L'estimation des composantes de la variance a été effectuée par maximum de vraisemblance restreint (REML). Les estimateurs ainsi obtenus présentent moins de biais que ceux qui seraient obtenus par maximum de vraisemblance (ML) (Pinheiro et Bates, 2000, Searle, 1995). De manière pratique, ces estimateurs sont obtenus directement en appliquant la fonction `lmer` du package `lme4`. L'objet obtenu par l'application de cette fonction est de classe `mer` et comporte une sortie `REmat` qui donne les composantes de la variance.

```
> res <- lmer(donnee[,k] ~ (1|Produit) + (1|Sujet) + (1|Sujet:Produit),
             data=donnee)
> varcomp <- summary(res)@REmat
```

4.2 La comparaison de variance

L'écart-type et la variance sont le reflet de la dispersion des données. Ils sont donc des indicateurs de la variabilité des données. Nous serons amenés à comparer la variabilité de jeux de données deux à deux (individuellement pour chaque variété). Pour ce faire, l'évolution des écart-types sera observée et testée par le test de comparaison de variance de Fisher-Snedecor.

Pour chaque variété i , la variance est calculée selon l'équation suivante :

$$s_i^2 = \frac{1}{(J * R) - 1} \sum_j \sum_r (X_{ijr} - \bar{X}_{i..})^2 \quad (\text{III-3})$$

où j correspond à l'indice des sujets et r à l'indice des répétitions.

Considérons une population de taille n , s^2 est un estimateur de σ^2 . Sous l'hypothèse normale,

$(n-1)\frac{s^2}{\sigma^2}$ suit une loi du Khi-deux à $(n-1)$ degrés de liberté.

Pour comparer deux échantillons de taille n_1 et n_2 , nous aurons $(n_1-1)\frac{s_1^2}{\sigma_1^2}$ et $(n_2-1)\frac{s_2^2}{\sigma_2^2}$.

Sous l'hypothèse $H_0: \sigma_1^2 = \sigma_2^2$, le rapport $\frac{\frac{n_1-1}{n_1-1} s_1^2}{\frac{n_2-1}{n_2-1} s_2^2}$ suit une loi de Fisher-Snedecor à $(n_1-1), (n_2-1)$

degrés de liberté.

En pratique, ce test est réalisé avec la fonction `var.test` du package `stats`. Pour comparer les variances, pour chaque descripteur k , de deux jeux de données nommées 'donnee1' et 'donnee2' comprenant K descripteurs, le code suivant est utilisé :

```
> res <- var.test (donnee1[,k], donnee2[,k], alternative="two.sided")
```

4.3 Corrélations

Les coefficients de corrélation (R) ont été utilisés. Ils ont été calculés par le test de corrélation de Pearson qui mesure l'intensité de la relation et l'efficacité de l'ajustement des données à un modèle linéaire.

En pratique, ce test est réalisé avec la fonction `cor` du package `stats`. Pour comparer le descripteur k et la mesure instrumentale 'instr' d'un jeu de données nommé 'donnee', le code suivant est utilisé :

```
> res <- cor (donnee[,k], donnee$instr, use="everything", method="pearson")
```

Partie IV. Résultats et Discussion

Chapitre I. Vérification de la sensibilité du panel

L'étude bibliographique a mis en avant la présence de variations perceptibles entre des fruits ayant différents stades de maturité et entre différentes parties d'un fruit. De plus, les fruits d'un lot sont susceptibles de présenter des différences de stade de maturité. L'hypothèse est donc que les sujets d'un panel entraînés sont suffisamment sensibles pour percevoir des différences entre des fruits d'un lot unique.

Ce premier chapitre de résultats est consacré à l'étude des capacités de discrimination du panel à deux niveaux : discrimination entre des lots correspondants à des variétés différentes, appelée *discrimination inter-variétale* ; et discrimination entre des sous-lots correspondants à des fruits d'une variété unique ayant des propriétés biologiques différentes, appelée *discrimination intra-variétale*.

La première partie de ce chapitre correspond à l'étude des capacités de discrimination inter-variétale du panel entre trois variétés – Ariane, Braeburn et Pink lady® – pour l'ensemble des descripteurs sensoriels. Les paramètres de performance d'accord et de répétabilité ont été observés afin de caractériser la qualité du panel. Par ailleurs, cette première partie permet de caractériser les trois variétés de pommes étudiées.

La deuxième partie présente les capacités de discrimination intra-variétale du panel. La capacité du panel à distinguer différents niveaux de fermeté acoustique a été testée et la hiérarchie des sous-lots de fermeté acoustique différente a été observée pour chaque variété. Trois niveaux de fermeté acoustique – bas, moyen et haut – ont été évalués pour chaque variété.

La troisième et dernière partie résume les capacités de discrimination du panel et fait état du possible impact de la sensibilité du panel sur la caractérisation de produits soumis aux variations biologiques que sont les pommes.

1. Étude de la capacité de discrimination inter-variétale du panel

1.1 Sélection des fruits

L'évaluation de la capacité de discrimination du panel ayant lieu au début du projet, et dans l'hypothèse d'un possible impact de la variabilité intra-lot sur la réponse sensorielle, une homogénéisation des fruits au sein de chaque variété a été réalisée par sonométrie.

Pour cette caractérisation, 400 fruits des trois variétés étudiées ont d'abord été évalués par sonométrie. Pour chaque variété, les fruits ont été classés selon leur fermeté acoustique S et les fruits ont été répartis en cinq groupes d'étendue égale (A à E par ordre de fermeté acoustique moyenne croissante) (Exemple pour la variété Ariane en Figure 10).

Les distributions des indices de fermeté acoustique sont présentées en Figure 11. Au sein d'une variété, les indices de fermeté acoustique sont variables, allant jusqu'à une différence de plus de $20 \text{ Hz}^2 \text{ g}^{2/3}$ entre les fruits extrêmes au sein des variétés Braeburn et Pink Lady®. De plus, les variétés Braeburn et Pink Lady® ont des caractéristiques acoustiques très similaires. Une ANOVA à un facteur variété fixe suivie d'une analyse post hoc confirme la proximité des variétés Braeburn et Pink Lady® pour l'indice de fermeté acoustique (Tableau 7).

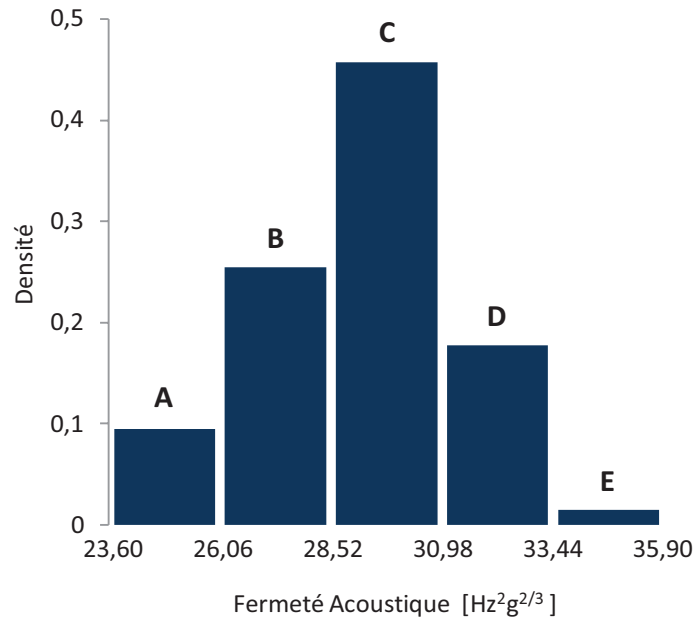


Figure 10. Caractérisation des fruits et répartition en cinq groupes (A à E) – exemple de la variété Ariane

Pour chaque variété, la fermeté acoustique moyenne est reportée dans le Tableau 7. Les 25 fruits dégustés par le panel ont été choisis dans le groupe C, dit de fermeté acoustique moyenne. L'intervalle de fermeté acoustique de ces fruits est également présenté dans le Tableau 7.

Lors de l'évaluation sensorielle, quatre répétitions ont été réalisées. Plusieurs sujets ont dégusté le même fruit – un à quatre sujets par fruit. Pour chaque répétition, l'ordre des variétés a été basé sur un carré latin de Williams. Les 19 sujets ont pris part à cette mesure.

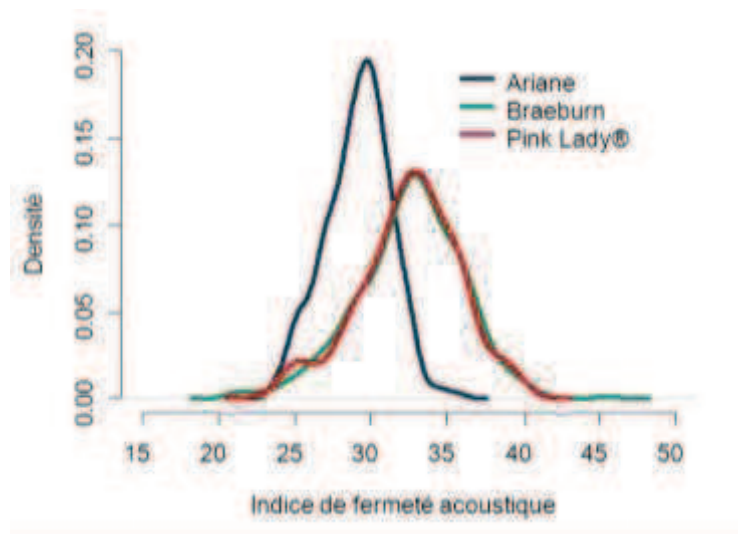


Figure 11. Performance inter-variétale – distribution de l'indice acoustique de fermeté pour les trois variétés

Tableau 7. Performance inter-variétale – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des fruits dégustés

Variété	Fermeté acoustique moyenne (Hz ² g ^{2/3})	Étendue de fermeté acoustique des fruits dégustés (Hz ² g ^{2/3})
Ariane	29,17 ± 2,15 a	[27,70 ; 31,10]
Braeburn	32,49 ± 3,61 b	[30,80 ; 32,30]
Pink Lady®	32,50 ± 3,29 b	[31,20 ; 32,50]

1.2 Analyse de la variance : modèle standard

Le modèle standard a été utilisé pour l'analyse des performances des sujets. Il est rappelé ci-après :

$$X_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijr} \quad (\text{IV-1})$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet*variété}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ représente l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction (aléatoire) sujet*variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Les trois indices de performance – la discrimination, l'accord et la répétabilité – ont été observés :

- La discrimination correspond au rapport du carré moyen de l'effet variété sur le carré moyen de l'interaction :

$$F_{\text{variété}} = \frac{CM_{\text{variété}}}{CM_{\text{sujet*variété}}} \quad (\text{IV-2})$$

- L'accord correspond au rapport du carré moyen de l'interaction sur le carré moyen de l'erreur :

$$F_{\text{sujet*variété}} = \frac{CM_{\text{sujet*variété}}}{CM_{\text{erreur}}} \quad (\text{IV-3})$$

- La répétabilité correspond à la racine carrée de la variance de l'erreur issue du modèle (IV-1) et représente l'écart moyen entre les répétitions pour chaque produit. Un indice élevé indique que le panel n'est pas répétable.

Le modèle (IV-1) a également été appliqué en considérant tous les effets comme aléatoires afin de calculer les contributions à la variance de chaque facteur. La variance totale du modèle aléatoire correspondant au modèle (IV-1) est donnée par l'équation (IV-4).

$$\text{Var}(X) = \sigma_{\text{variété}}^2 + \sigma_{\text{sujet}}^2 + \sigma_{\text{sujete*variété}}^2 + \sigma_e^2 \quad (\text{IV-4})$$

1.3 Résultats & discussion

Les résultats du modèle (IV-1) pour les sept descripteurs sont présentés dans le Tableau 8. Les résultats indiquent un effet variété significatif pour tous les descripteurs. Le panel est donc capable de détecter des différences sensorielles entre les trois variétés présentées.

Les résultats du test post-hoc permettent d'obtenir plus de précisions sur les variétés perçues différentes. Les notes moyennes de chaque variété pour chaque descripteur et les résultats du test post hoc sont présentés dans le Tableau 9. Les profils des trois variétés sont différents bien que les niveaux de fermeté acoustique observés soient proches, notamment pour les variétés Braeburn et Pink Lady® (allant de 30,8 Hz²g^{2/3} à 32,3 Hz²g^{2/3} et 31,2 Hz²g^{2/3} à 32,5 Hz²g^{2/3}, respectivement). Ce résultat peut s'expliquer par une dépendance variétale de la mesure acoustique (Molina-Delgado *et al.*, 2009b). Les trois variétés sont bien différenciées pour les descripteurs de texture *croquant*, *crissant* et *jutosité*. La variété Pink Lady® est la plus croquante (8,02), crissante (7,45) et juteuse (7,33). Au contraire, la variété Ariane est la moins croquante (5,45), crissante (4,75) et juteuse (5,07). La variété Braeburn est intermédiaire mais présente des caractéristiques plus proches de la variété Ariane, tout en restant significativement différente (*croquant* : 6,00 ; *crissant* : 5,75 ; *jutosité* : 5,93). Concernant les descripteurs *fermeté* et *fondant*, deux groupes se distinguent. Les variétés Ariane et Braeburn présentent des similarités pour ces caractéristiques et sont les moins fermes (respectivement 4,22 et 4,49) et les plus fondantes (respectivement 5,06 et 4,41). La variété Pink Lady® présente la note la plus élevée pour le descripteur *fermeté* (7,51) et la note la plus basse pour le descripteur *fondant* (1,63). Concernant les saveurs, les trois variétés sont différenciées pour le descripteur *acidité*. La variété Ariane est la plus acide (5,46). La variété Braeburn est la moins acide (3,78). La variété Pink lady® est intermédiaire bien que présentant une acidité (3,95) plus proche de celle de la variété Braeburn que de la variété Ariane. Les notes moyennes pour le descripteur *sucré* sont proches. La variété Ariane est la plus sucrée (4,96) et la variété Braeburn la moins sucrée (4,33). La variété Braeburn est la moins sapide avec les notes les plus faibles pour les descripteurs *acidité* et *sucré*.

La description des variétés faite par le panel est proche de la description de ces variétés donnée par Vaysse et Landry (2004). Les descripteurs cités dans cet ouvrage sont : *croquant*, *jutosité*, *farinosité*, *couleur rouge*, *parfum*, *goût*, *note fruits rouges*, *acidité*, *sucré*. La variété Pink lady® présente les notes les plus élevés pour les descripteurs *croquant* et *jutosité* et la note la plus basse pour le descripteur *farinosité*. Au contraire, les variétés Ariane et Braeburn sont décrites comme étant assez similaires et présentent les notes les plus basses pour les descripteurs *croquant* et *jutosité* et les notes les plus élevées pour le descripteur *farinosité*. Les descriptions en termes de saveurs sont concordantes.

Les résultats d'analyse de la variance (Tableau 8) indiquent également un effet sujet significatif pour tous les descripteurs comme dans la plupart des études sensorielles (Pineau, 2006). Il n'apporte pas réellement d'information sur les performances des sujets et indique uniquement que le niveau moyen de notation des sujets est variable. Le terme d'interaction sujet*variété n'est significatif pour aucun descripteur, preuve de l'absence de désaccord entre les sujets. Concernant l'indice de répétabilité (σ^2), il varie de 1,40 à 2,08 en fonction du descripteur. Cet indice a été reporté pour 29 descripteurs évalués sur une échelle continue (0 à 10) par un panel de 11 sujets pour la description de neuf jus de mandarine (Carbonell *et al.*, 2007) sans être toutefois utilisé pour évaluer la répétabilité du panel mais seulement la contribution de chaque sujet à la non répétabilité du panel. Les valeurs indiquées varient de 0,62 à 3,39 en fonction du descripteur. Les auteurs ont conclu à une performance satisfaisante du panel. L'observation de ces valeurs permet de conclure à une relative bonne répétabilité de notre panel.

Tableau 8. Performance inter-variétale – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les sept descripteurs

Descripteurs	Variété		Sujet		Sujet*variété		Erreur σ^2
	$F_{2; 207}$	p-value	Chi ²	p-value	Chi ²	p-value	
Croquant	53,39	2,01E-19***	18,20	1,99E-05***	0,00	1,00E+00	1,61
Fermeté	92,96	1,56E-29***	42,33	7,72E-11***	0,00	1,00E+00	1,65
Crissant	67,63	2,50E-23***	73,33	1.10E-17***	0,86	3,50E-01	1,40
Jutosité	28,96	8,14E-12***	22,09	2,61E-06***	1,58	2,10E-01	1,78
Fondant	57,90	1,07E-20***	31,94	1,59E-08***	0,00	1,00E+00	2,08
Acidité	22,27	1,74E-09***	93,22	4.68E-22***	0,00	9,90E-01	1,72
Sucre	3,26	4,04E-02*	87,24	9.63E-21***	0,30	5,87E-01	1,49

Niveau de significativité : 0,05* ; 0,01** ; 0,001***

Tableau 9. Performance inter-variétale – notes moyennes et erreurs standards pour les trois variétés de pommes et les sept descripteurs

Descripteurs	Ariane	Braeburn	Pink Lady®
Croquant	5,45 ± 0,25 a ¹	6,00 ± 0,25 b	8,02 ± 0,25 c
Fermeté	4,22 ± 0,31 a	4,49 ± 0,31 a	7,51 ± 0,31 b
Crissant	4,75 ± 0,32 a	5,75 ± 0,32 b	7,45 ± 0,32 c
Jutosité	5,07 ± 0,30 a	5,93 ± 0,30 b	7,33 ± 0,30 c
Fondant	5,06 ± 0,36 b	4,41 ± 0,36 b	1,63 ± 0,36 a
Acidité	5,46 ± 0,42 c	3,78 ± 0,42 a	3,95 ± 0,42 b
Sucré	4,96 ± 0,36 b	4,33 ± 0,36 a	4,57 ± 0,36 ab

¹ Les valeurs suivies de la même lettre minuscule sur une même ligne ne sont pas différentes au seuil de 5%

L'observation des contributions à la variance (Tableau 10 et Figure 12) indique que le facteur variété contribue fortement à la variabilité de la réponse sensorielle (23 à 46%) sauf pour les descripteurs *acidité* et *sucré* (13% et 1% respectivement). En général, la contribution du facteur sujet est inférieure à la contribution du facteur variété (là encore excepté pour les descripteurs *acidité* et *sucré*). Quant au terme d'interaction sujet*variété, quel que soit le descripteur, sa contribution est négligeable (<6%). Néanmoins, la part de variance non-expliquée (erreur), et attribuée à la moindre répétabilité des sujets, est importante (représentant plus de 37% de la variabilité totale). La forte contribution de l'erreur résiduelle à la variance pourrait être expliquée par des facteurs non présents dans l'analyse comme par exemple la variabilité intra-lot.

Tableau 10. Performance inter-variétale – composantes de la variance et, entre parenthèses, leur pourcentage de la variation totale pour les sept descripteurs dans le cas du modèle (IV-1) aléatoire

Descripteurs	Variété	Sujet	Sujet*Variété	Erreur
Croquant	1,79 (36)	0,53 (11)	0,00 (00)	2,59 (53)
Fermeté	3,30 (46)	1,11 (16)	0,00 (00)	2,73 (38)
Crissant	1,82 (34)	1,38 (26)	0,17 (03)	1,97 (37)
Jutosité	1,24 (23)	0,72 (13)	0,35 (06)	3,16 (58)
Fondant	3,25 (36)	1,38 (15)	0,05 (01)	4,31 (48)
Acidité	0,82 (13)	2,58 (41)	4,73 (00)	2,95 (46)
Sucré	0,06 (01)	1,84 (43)	0,12 (03)	2,22 (53)

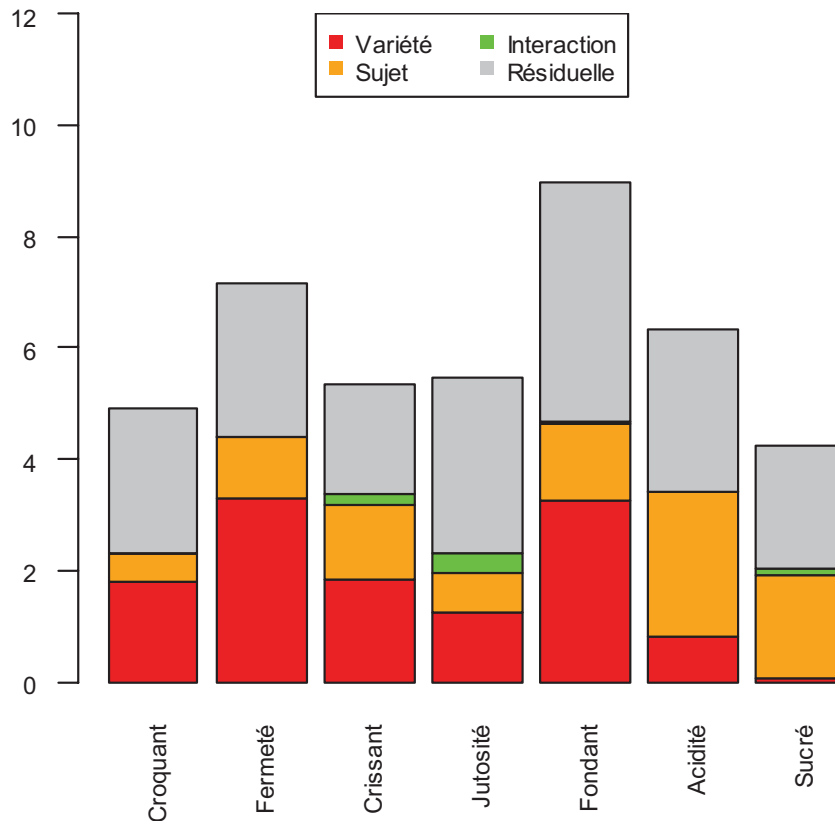


Figure 12. Performance inter-variétale – estimations des composantes de la variance pour les sept descripteurs

1.4 Conclusion

Pour conclure, le panel est capable de différencier les trois variétés de pommes étudiées pour les descripteurs *croquant*, *crissant*, *jutosité* et *acidité*. Le panel distingue également la variété Pink lady® des variétés Ariane et Braeburn pour les descripteurs *fermeté* et *fondant*. Concernant le descripteur *sucré*, il est perçu différent entre les variétés Ariane et Braeburn. Des différences de notations sont observées. Les sujets n'ont pas la même moyenne de notation. De plus, les sujets sont en accord.

Cependant, une part importante des variations de la réponse sensorielle reste inexpliquée (37 à 58% en fonction du descripteur) et est attribuée généralement à la non-répétabilité du panel. La variabilité des fruits a été suggérée comme cause de la variabilité totale de la réponse à de nombreuses reprises dans les études sensorielles (e.g. Hampson *et al.* (2000) ; Mehinagic *et al.* (2003) ; Seppä *et al.* (2012) ; Williams A.A. et Carter (1977)). Nous supposons alors qu'une partie de la variabilité non expliquée pourrait l'être par les variations intra-lot des fruits.

Dans la suite de ce chapitre, les capacités du panel entraîné à discriminer les fruits d'un lot unique d'une variété donnée, identifiés comme présentant des variations biologiques, sont étudiées.

2. Étude la capacité de discrimination intra-variétale du panel

2.1 *Sélection des fruits*

Pour cette mesure, environ 480 fruits des trois variétés étudiées ont d'abord été caractérisés par sonométrie. 30% de ces fruits avaient préalablement été conservés 3 semaines à 18°C pour augmenter les différences biologiques au sein de chaque variété (lot). Pour chaque variété, les fruits ont été classés selon leur fermeté acoustique *S* et ont été répartis en cinq groupes d'étendue égale (A à E par ordre de fermeté acoustique moyenne croissante). Ces groupes étaient différents de ceux déterminés pour la mesure de discrimination inter-variétale.

Les distributions des indices de fermeté acoustique des lots initiaux sont présentées en Figure 13. Les distributions sont bimodales, ce qui correspond à la conséquence des deux conditions de stockage (4°C et 18°C). Cette expérimentation ayant été réalisée trois semaines après l'expérimentation sur la capacité de discrimination inter-variétale, une chute de fermeté acoustique est également observée pour les fruits stockés à 4°C, notamment pour la variété Braeburn.

Afin d'évaluer la capacité de discrimination du panel entre des sous-lots d'une même variété présentant des niveaux de fermeté acoustique différents, pour chaque variété, trois sous-lots de niveau de fermeté acoustique différents ont été constitués. Les fruits ont été sélectionnés dans les groupes A, C et E, dits de fermeté acoustique basse, moyenne et haute. Neuf sous-lots (trois variétés*trois niveaux de fermeté acoustique) ont donc été déterminés. Pour chacun de ces sous-lots, 18 fruits ont été sélectionnés et dégustés par un à trois sujets au cours de deux répétitions. Pour chaque répétition, l'ordre des sous-lots a été basé sur un carré latin de Williams. 18 sujets ont pris part à cette mesure.

Pour chaque variété (lot), la fermeté acoustique moyenne est reportée dans le Tableau 11, ainsi que l'intervalle de fermeté acoustique de chaque sous-lot.

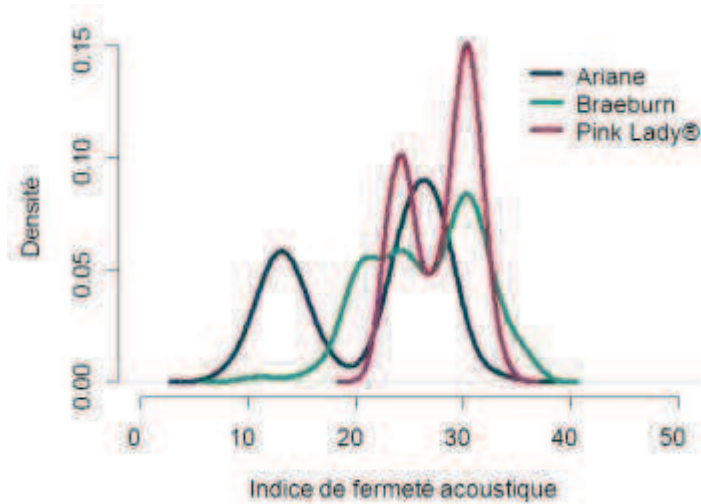


Figure 13. Performance intra-variétale – distribution de l'indice acoustique de fermeté pour les trois variétés

Tableau 11. Performance intra-variétale – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des trois sous-lots dégustés pour chaque variété

Variété	Fermeté acoustique moyenne	Étendue de fermeté acoustique des fruits dégustés		
		Niveau Bas	Niveau Moyen	Niveau Haut
Ariane	21,28 ± 6,76	[9,60 ; 11,50]	[17,30 ; 23,00]	[28,60 ; 29,50]
Braeburn	26,32 ± 5,46	[15,10 ; 18,90]	[25,30 ; 26,10]	[33,60 ; 35,40]
Pink Lady®	27,95 ± 3,20	[22,10 ; 22,80]	[27,40 ; 28,10]	[31,70 ; 32,00]

2.2 Analyse de la variance : modèle mixte

Dans le cas présent, le facteur produit a pu être décomposé en deux facteurs : variété et fermeté acoustique. Afin de vérifier la capacité de discrimination du panel, un modèle d'ANOVA à facteurs croisés a été appliqué, avec un facteur variété en effet fixe, un facteur fermeté acoustique en effet fixe, ces deux facteurs étant croisés et considérés comme qualitatifs. Le facteur sujet (aléatoire) a été ajouté. Enfin l'interaction variété*fermeté acoustique*sujet (aléatoire) a été ajoutée au modèle.

$$Y_{ifjr} = \mu + \nu_i + \psi_f + \lambda_{if} + \alpha_j + \gamma_{ifj} + \varepsilon_{ifjr} \quad (\text{IV-5})$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ifj} \sim N(0, \sigma_{\text{sujet*variété*fermeté acoustique}}^2)$$

$$\varepsilon_{ifjr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ représente l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), ψ_f à l'effet fixe fermeté acoustique, λ_{if} à l'interaction (fixe) variété*fermeté acoustique, α_j à l'effet aléatoire sujet, γ_{ijf} à l'interaction (aléatoire) sujet*variété*fermeté acoustique et ε_{ijfr} à l'erreur résiduelle (aléatoire) du modèle.

Les indices observés étaient :

- La discrimination variétale et la discrimination des niveaux de fermeté acoustique, qui correspond au rapport du carré moyen de l'effet produit ou fermeté acoustique sur le carré moyen de l'interaction :

$$F_{\text{variété}} = \frac{CM_{\text{variété}}}{CM_{\text{sujet*variété*fermeté acoustique}}} \quad (\text{IV-6})$$

$$\text{et } F_{\text{fermeté acoustique}} = \frac{CM_{\text{fermeté acoustique}}}{CM_{\text{sujet*variété*fermeté acoustique}}} \quad (\text{IV-7})$$

Par ailleurs la statistique du terme d'interaction variété-par-fermeté-acoustique est analysée afin de savoir si la discrimination des niveaux de fermeté acoustique dépend de la variété.

$$F_{\text{variété*fermeté acoustique}} = \frac{CM_{\text{variété*fermeté acoustique}}}{CM_{\text{sujet*variété*fermeté acoustique}}} \quad (\text{IV-8})$$

- L'accord correspond au rapport du carré moyen de l'interaction sur le carré moyen de l'erreur :

$$F_{\text{sujet*variété*fermeté acoustique}} = \frac{CM_{\text{sujet*variété*fermeté acoustique}}}{CM_{\text{erreur}}} \quad (\text{IV-9})$$

- La répétabilité correspond à la racine carrée de la variance de l'erreur issue du modèle (IV-4-1) et représente l'écart moyen entre les répétitions pour chaque produit. Un indice élevé indique que le panel n'est pas répétable.

2.3 Résultats & discussion

En plus d'être capable de différencier des fruits appartenant à différentes variétés (lots), le panel est capable de différencier des fruits au sein d'une même variété (même lot) avec des niveaux de fermeté acoustique différents. Les résultats d'ANOVA présentés dans le Tableau 12 montrent non seulement un effet significatif du facteur variété pour les sept descripteurs, mais aussi un effet significatif du facteur fermeté acoustique pour les descripteurs *croquant*, *fermeté*, *jutosité* et *acidité*.

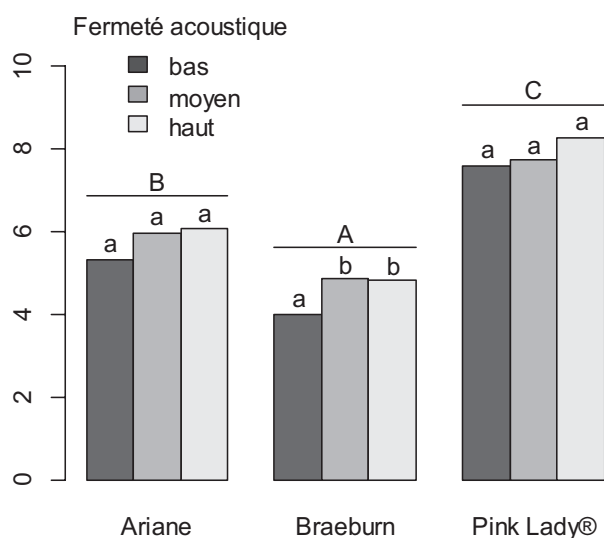
De plus, la perception de différences entre les niveaux de fermeté acoustique dépend de la variété pour le descripteur *sucré* ($F_{\text{variété*fermeté acoustique}}$ significative).

La capacité de discrimination inter-variétale est confirmée. Cependant l'analyse post hoc (Tableau 13) indique des différences comparées aux résultats obtenus lors de discrimination inter-variétale. Dans le cas présent, concernant la texture, les trois variétés sont bien différenciées pour les descripteurs *croquant*, *fermeté*, *crissant* et *fondant* et les variétés Ariane et Braeburn sont perçues proches pour le descripteur *jutosité*. Conformément aux résultats précédents, la Pink Lady® obtient les notes les plus élevées pour les descripteurs *croquant* (7,86), *fermeté* (7,37), *crissant* (7,16) et *jutosité* (7,27), et la note la plus basse pour le descripteur *fondant* (1,87). Contrairement aux résultats précédents, qui indiquaient de plus faibles notes pour la variété Ariane comparée à la variété Braeburn pour les descripteurs *croquant*, *crissant* et *jutosité*, ici la variété Braeburn obtient les notes les plus basses ($4,56 < 5,78$; $4,53 < 5,39$ et $5,21 < 4,93$ respectivement pour *croquant*, *crissant* et *jutosité*). De plus, pour les descripteurs *fermeté* et *fondant*, ces deux variétés n'étaient pas différenciées et le sont maintenant avec la variété Braeburn perçue moins ferme ($3,12 < 4,92$) et plus fondante ($6,81 > 4,40$) que la variété Ariane. Concernant les saveurs, les résultats concordent avec ceux obtenus précédemment pour le descripteur *acidité* et les trois variétés sont non différenciées pour le descripteur *sucré*.

Les différences observées concernant le classement des variétés, et notamment des variétés Ariane et Braeburn, s'expliquent par les différences d'évolution de ces fruits au cours du stockage. En effet, les pommes de la variété Braeburn stockées à 4°C semblent avoir plus évolué que les pommes des variétés Ariane et Pink Lady®. Sur la Figure 13, contrairement aux distributions des variétés Ariane et Pink Lady®, la distribution de la variété Braeburn présente une distribution bimodale moins marquée. L'observation de l'évolution des fermetés acoustiques des fruits mesurés lors des deux expérimentations montre une plus forte chute des valeurs pour la variété Braeburn (données non présentées).

Concernant les différences intra-variétale entre les différents niveaux de fermeté acoustique, l'analyse post hoc (Tableau 13) montre que pour la variété Ariane, des différences entre les niveaux de fermeté acoustique sont observées pour les descripteurs *jutosité* et *acidité*. Concernant le descripteur *jutosité*, les fruits des niveaux de fermeté acoustique bas et moyen ne sont pas différenciés mais obtiennent des notes significativement inférieures au niveau haut ($[4,36 ; 4,75] < 5,68$). Les acidités perçues sont significativement différentes et sont croissantes du niveau bas au niveau haut ($3,39 < 4,33 < 5,44$). Pour la variété Braeburn, les fruits aux niveaux de fermeté

acoustique différents sont différenciés pour les descripteurs *croquant*, *jutosité*, *acidité* et *sucré*. Pour les descripteurs de texture, les fruits ayant des niveaux de fermeté acoustique bas se distinguent des fruits ayant des niveaux moyens et haut par les descripteurs *croquant* et *jutosité* et présentent des notes inférieures ($3,99 < [4,81 ; 4,87]$ et $4,09 < [5,72 ; 5,81]$ pour *croquant* et *jutosité* respectivement). Concernant le descripteur *acidité*, les fruits de niveau de fermeté acoustique bas sont perçus moins acides que les fruits de niveau moyen ($2,74 < 3,57$). Les fruits de niveau de fermeté acoustique haut obtiennent une note intermédiaire mais non différente. Concernant la perception du descripteur *sucré*, elle est moindre pour les fruits de niveaux de fermeté acoustique bas et moyen (non différenciés) ($[4,26 ; 4,52] < 5,35$). Quant à la variété Pink Lady®, seules des différences pour le descripteur *acidité* sont perçues par le panel avec une acidité inférieure pour les fruits de niveau de fermeté acoustique bas ($2,61 < [3,64 ; 4,26]$). Par ailleurs, des tendances de discrimination ($0,05 < p\text{-value} < 0,10$) sont observées pour la variété Ariane et le descripteur *croquant*, la variété Braeburn et les descripteurs *fermeté* et *fondant*, et la variété Pink Lady® et les descripteurs *croquant*, *fermeté*, *jutosité* et *sucré*. Pour illustrer ces tendances, la Figure 14 représente les différences au sein des variétés et entre les variétés pour le descripteur *croquant*.



Les lettres minuscules correspondent aux différences entre niveaux de fermeté acoustique au sein de chaque variété avec un seuil 5%,
Les lettres majuscules correspondent aux différences entre les notes moyennes des variétés avec un seuil 5%

Figure 14. Notes moyennes pour le descripteur *croquant*, les trois variétés et les trois niveaux de fermeté acoustique

Il est à noter que, quelle que soit la variété, pour les descripteurs *croquant*, *fermeté*, *crissant* et *jutosité*, et que les différences perçues soient significatives ou non, le niveau de fermeté acoustique

bas obtient la note la plus faible (excepté pour Pink Lady®/*crissant*). Pour le descripteur *fondant*, les pommes les moins fermes d'après la mesure acoustique (niveau bas) obtiennent des notes plus faibles que les pommes de niveau de fermeté acoustique moyen et haut. Des travaux (Mehinagic *et al.*, 2006, Molina-Delgado *et al.*, 2009b) ont montré que la mesure acoustique discriminait les fruits selon leur maturité et que les pommes plus mûres avaient des fermetés acoustiques inférieures. De plus, les pommes ayant des niveaux de maturité plus faibles sont caractérisées par des notes inférieures pour les descripteurs de texture tels que *croquant*, *fermeté* et *jutosité* et supérieure pour le descripteur *fondant* (Harker *et al.*, 2002b, Mehinagic, 2004, Mehinagic *et al.*, 2006, Seppä *et al.*, 2012, Watada *et al.*, 1980). Nos résultats vont dans ce sens, les sous-lots ayant une fermeté acoustique plus faible étant moins croquants et moins fermes (tendance). Concernant les saveurs, les pommes avec un niveau de fermeté acoustique bas sont perçues comme moins acides quelle que soit la variété. Il a été démontré que des pommes plus mûres présentent des teneurs plus faibles en acide (Ackermann *et al.*, 1992, Plotto *et al.*, 1999). L'observation de notes pour le descripteur *acidité* inférieures pour les pommes de fermeté acoustique inférieure semble donc cohérente et concorde avec l'observation d'une diminution de l'acidité quand la maturité augmente par Mehinagic *et al.* (2004). Les résultats de discrimination intra-variétale concordent également avec la capacité démontrée d'un panel à discriminer des niveaux de maturité différents (Harker *et al.*, 2002c, Seppä *et al.*, 2012, Watada *et al.*, 1980).

L'interaction sujet*variété*fermeté acoustique n'est pas significative, excepté pour le descripteur *jutosité* avec un effet observé au seuil de 5%. Les sujets sont donc en accord excepté pour le descripteur *jutosité* (p-value = 4,27E-02*). Le niveau de l'indice de répétabilité varie de 1,47 à 1,73 en fonction du descripteur. Le niveau de cet indice est proche du niveau observé pour la performance inter-variété.

Tableau 12. Performance intra-variétale – effet des facteurs variété, fermeté acoustique, sujet et des interactions variété*fermeté acoustique et sujet*variété*fermeté acoustique sur les notes des neuf produits (trois variétés de pomme*trois fermeté acoustique) et les sept descripteurs

Descripteur	Variété		Fermeté Acoustique		Variété*Fermeté Acoustique		Sujet		Sujet*Variété* Fermeté Acoustique		Erreur σ^2
	F _{2;302}	p-value	F _{2;302}	p-value	F _{4;298}	p-value	Chi ²	p-value	Chi ²	p-value	
Croquant	121,30	2.16E-39***	6,43	1.84E-03***	0,60	6.63E-01	49,98	1.56E-12***	0,00	1.00E+00	1,57
Fermeté	186,49	1.82E-53***	3,65	2.71E-02*	0,66	6.20E-01	52,99	3.35E-13***	0,00	1.00E+00	1,62
Crissant	80,51	8.74E-29***	1,02	3.62E-01	0,88	4.76E-01	50,14	1.43E-12***	0,88	3.49E-01	1,47
Jutosité	51,01	2.29E-17***	12,39	1.11E-05***	1,82	1.28E-01	22,15	2.53E-06***	4,11	4.27E-02*	1,51
Fondant	196,19	1.93E-55***	1,20	3.03E-01	0,57	6.85E-01	74,70	5.47E-18 ***	1,28	2.57E-01	1,73
Acidité	11,98	9.85E-06***	20,23	5.69E-09***	1,46	2.14E-01	134,53	4.19E-31 ***	3,31	6.89E-02	1,60
Sucré	4,62	1.06E-02*	0,19	8.27E-01	3,37	1.02E-02*	119,56	7.89E-28 ***	0,00	9.98E-01	1,55

Niveau de significativité : 0,05* ; 0,01** , 0,001***

Tableau 13. Performance intra-variétale - notes moyennes pour les trois variétés de pomme, les trois niveaux de fermeté acoustique et les sept descripteurs

Descripteur	Ariane				Braeburn				Pink Lady®			
	Fermeté acoustique			Moyenne variété	Fermeté acoustique			Moyenne variété	Fermeté acoustique			Moyenne variété
	Bas	Moyen	Haut		Bas	Moyen	Haut		Bas	Moyen	Haut	
Croquant	5,33 a ¹ (a) ²	5,96 a(b)	6,05 a(b)	5,78 B ³	3,99 a	4,87 b	4,81 b	4,56 A	7,59 a(a)	7,72 a(ab)	8,27 a(b)	7,86 C
Fermeté	4,68 a	4,80 a	5,28 a	4,92 B	2,73 a(a)	3,39 a(b)	3,24 a(ab)	3,12 A	7,13 a(a)	7,18 a(a)	7,81 a(b)	7,37 C
Crissant	5,17 a	5,40 a	5,59 a	5,39 B	4,32 a	4,86 a	4,41 a	4,53 A	7,08 a	6,97 a	7,42 a	7,16 C
Jutosité	4,36 a	4,75 a	5,68 b	4,93 A	4,09 a	5,81 b	5,72 b	5,21 A	6,88 a(a)	7,29 a(ab)	7,65 a(b)	7,27 B
Fondant	4,39 a	4,60 a	4,20 a	4,40 B	7,27 a(b)	6,69 a(ab)	6,46 a(a)	6,81 C	1,94 a	1,88 a	1,78 a	1,87 A
Acidité	3,39 a	4,33 b	5,44 c	4,39 C	2,74 a	3,57 b	3,54 ab	3,28 A	2,61 a	3,64 b	4,26 b	3,50 B
Sucré	5,04 a	5,23 a	4,99 a	5,09 A	4,52 a	4,26 a	5,35 b	4,71 A	5,38 a(ab)	5,68 a(b)	4,98 a(a)	5,35 A

¹ Les valeurs suivies de la même lettre minuscule au sein d'une variété ne sont pas différentes au seuil de 5%

² Les valeurs suivies de la même lettre minuscule entre parenthèses au sein d'une variété ne sont pas différentes au seuil de 10%

³ Les valeurs suivies de la même lettre majuscule sur une même ligne ne sont pas différentes au seuil de 5%

2.4 Conclusion

Pour conclure, pour certaines des variétés et certains des descripteurs étudiés, le panel est capable de percevoir des différences entre les fruits avec des niveaux de fermeté acoustique différents. La discrimination des niveaux de fermeté acoustique est dépendante de la variété. Par exemple, les différences perçues entre les niveaux de fermeté acoustique pour la variété Braeburn sont plus importantes que pour les variétés Ariane et Pink Lady®. Cette observation indique non seulement une différence de variabilité entre les variétés mais aussi une dépendance de la mesure acoustique par rapport aux variétés. En effet, la meilleure discrimination entre les niveaux de fermeté acoustique observée pour la variété Braeburn peut s'expliquer par des différences supérieures entre les niveaux de fermeté ($\Delta = 14,6 \text{ Hz}^2\text{g}^{2/3}$) par rapport aux différences observées pour la variété Pink Lady® ($\Delta = 8,8 \text{ Hz}^2\text{g}^{2/3}$). Cette hypothèse n'est cependant pas valable lorsque l'on compare les différences au sein de la variété Braeburn avec celles observées pour la variété Ariane ($\Delta = 15,6 \text{ Hz}^2\text{g}^{2/3}$). Dans ce cas, il serait possible d'expliquer la meilleure discrimination observée pour la variété Braeburn par une dépendance variétale de la mesure acoustique, déjà observée par Molina-Delgado *et al.* (2009a). Des différences acoustiques de même amplitude ne correspondent pas à une perception similaire de différence sensorielle en fonction de la variété. Il a également été démontré que la mesure de fermeté acoustique dépendait de la forme des fruits (Chen et DeBaerdemaeker, 1993), or les fruits de variétés Ariane, Braeburn et Pink Lady® sont de formes différentes.

Même si dans le cas présent le stockage des fruits à une température supérieure pendant quelques semaines a augmenté la variabilité biologique, les différences observées ne semblent pas exagérées. Cette condition de stockage à température ambiante peut être considérée comme une simulation des conditions observées dans les surfaces commerciale (Róth *et al.*, 2007). Dans ces travaux, des différences de l'ordre de $10 \text{ Hz}^2\text{g}^{2/3}$ ont été observées entre les fruits à la récolte et après 14 jours de conservation à température ambiante. Les valeurs de fermeté acoustique observées dans notre cas ne semblent donc pas aberrantes comparées aux références antérieures.

3. Conclusion générale

L'étude des capacités de discrimination du panel a clairement démontré la bonne sensibilité du panel et sa faculté à percevoir des différences non seulement entre des fruits de variétés (lots) différentes

mais aussi entre les fruits d'une même variété (même lot) avec des caractéristiques différentes (fermeté acoustique). La discrimination intra-variétale dépend cependant des variétés et des descripteurs.

Par ailleurs, dans le cadre de l'étude de discrimination inter-variétale, une part importante de variation non expliquée a été observée. Cette erreur est généralement attribuée à la moindre répétabilité du panel. Cependant, la sensibilité du panel à de faibles différences entre les fruits au sein de chaque variété (lot) pourrait expliquer une part de cette erreur. Il serait donc intéressant de tenir compte des différences existant entre les fruits d'un lot et perçues par le panel dans l'analyse de la réponse sensorielle. La prise en compte de la variabilité biologique nous semble essentielle afin de conclure de façon fiable sur les différences existant entre des lots.

Chapitre II. Mesurer la variabilité : sélection d'un modèle adapté

À l'issue des résultats présentés au Chapitre I, ce deuxième chapitre est consacré à l'adaptation du modèle d'analyse de la variance pour le traitement des données sensorielles issues de la caractérisation de pommes par la méthode du profil conventionnel.

La première partie de ce chapitre explore le potentiel d'un modèle d'analyse de la variance dans lequel un facteur permettant de prendre en compte la variabilité intra-lot biologique, à savoir un effet fruit, est ajouté. Ce modèle sera appelé *modèle « fruit »*. L'importance du facteur fruit est évaluée et les résultats de ce modèle sont comparés à ceux issus du modèle standard.

La deuxième partie de ce chapitre correspond à l'application d'une approche basée sur le modèle « assessor » proposé par Brockhoff et Skovgaard (1994), en complément du modèle « fruit ». Dans un premier temps, la nature des différences existant entre les sujets du panel est observée. Puis, l'impact de la distinction entre effet dispersion et désaccord dans le traitement des données est évalué.

La troisième partie présente un bilan des apports des deux modèles étudiés, dans le cadre de l'évaluation de pommes.

1. Étude du modèle « fruit » incluant la prise en compte de la variabilité intra-lot

La capacité du panel entraîné à discriminer des pommes d'une même variété (un lot) mais présentant des différences biologiques a été démontrée. L'hétérogénéité d'un lot de pommes peut

donc être perçue par le panel. Ajoutés à l'étude bibliographique, ces résultats confirment le besoin d'adapter le traitement des données sensorielles issues du profil de pommes en prenant en compte leur variabilité biologique. Afin d'évaluer les implications de l'ajout d'un facteur 'fruit' permettant de quantifier la variabilité intra-lot dans le modèle d'analyse de la variance, nous reprenons les données issues de la mesure de discrimination inter-variétale (présentées au Chapitre I.1). Pour aller plus loin, des données sont simulées afin de mettre en exergue les conséquences de la non-prise en compte de la variabilité des fruits.

1.1 Les modèles

1.1.1 Le modèle standard

Le modèle standard correspond au modèle le plus couramment utilisé pour analyser la réponse sensorielle. Il est rappelé ci-après (IV-10).

$$X_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijr} \quad (\text{IV-10})$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet*variété}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ représente l'effet moyen global, ν_i correspond à l'effet fixe variété, α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction (aléatoire) sujet*variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Il est utile de rappeler que l'interaction est souvent interprétée comme le désaccord entre les sujets et que le terme d'erreur est souvent utilisé comme indicateur de répétabilité.

1.1.2 Le modèle « fruit » tenant compte de la variabilité intra-lot

Les différences effectives entre les fruits d'un lot peuvent être suffisamment importantes pour être perçues par un panel entraîné. Or, l'utilisation du modèle standard ne tient pas compte de l'éventuelle existence de ces différences. Le désaccord des sujets et le manque de répétabilité du panel pourrait cependant être dus à ces différences entre les fruits. Une adaptation du modèle standard est proposée. Cette adaptation consiste en l'ajout d'un facteur permettant de prendre en compte la variabilité intra-lot. Le modèle proposé est le suivant et sera appelé *modèle « fruit »* :

$$X_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varphi_{f(ijr)} + \varepsilon_{ijr} \quad (\text{IV-11})$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet*variété}}^2)$$

$$\varphi_{f(ijr)} \sim N(0, \sigma_{\text{fruit(variété)}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ correspond à l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction (aléatoire) sujet*variété, $\varphi_{f(ijr)}$ à l'effet aléatoire fruit imbriqué dans le facteur variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

L'ajout du facteur fruit implique des changements quant au calcul des carrés moyens attendus et donc des tests de significativité des effets. Le Tableau 14 contient les carrés moyens attendus et les statistiques qui en résultent. L'ajout de l'effet fruit aléatoire au modèle modifie la statistique de test de l'effet variété. Le carré moyen de l'effet variété est comparé à une combinaison du carré moyen de l'interaction sujet*variété et du carré moyen de l'effet fruit. Un effet variété significatif indique donc que les différences entre variétés sont plus importantes que la somme du désaccord global des sujets et des différences entre les fruits au sein de chaque variété. Dans le cas d'une forte variabilité biologique, il est judicieux de comparer l'importance des différences entre les variétés à l'importance des différences au sein des variétés.

Tableau 14. Test des effets des facteurs du modèle « fruit »

Effet	Carré moyen attendu	Statistique
Variété	$JR \sigma_{\text{variété}}^2 + R \sigma_{\text{sujet*variété}}^2 + P \sigma_{\text{fruit}}^2 + \sigma_e^2$	$\frac{CM_{\text{variété}}}{CM_{\text{sujet*variété}} + CM_{\text{fruit}} - CM_{\text{erreur}}}$
Sujet	$IR \sigma_{\text{sujet}}^2 + R \sigma_{\text{sujet*variété}}^2 + \sigma_e^2$	$\frac{CM_{\text{sujet}}}{CM_{\text{sujet*variété}}}$
Interaction	$R \sigma_{\text{sujet*variété}}^2 + \sigma_e^2$	$\frac{CM_{\text{sujet*variété}}}{CM_{\text{erreur}}}$
Fruit(variété)	$P \sigma_{\text{fruit}}^2 + \sigma_e^2$	$\frac{CM_{\text{fruit}}}{CM_{\text{erreur}}}$
Erreur	σ_e^2	

L'apport de ce nouveau modèle a été évalué en observant dans un premier temps la significativité et l'importance du facteur fruit dans l'explication de la variabilité des résultats. L'évaluation de l'importance des différents facteurs dans l'explication de la variabilité totale des résultats a été effectuée en appliquant le modèle (IV-11) en considérant tous les effets comme aléatoires. La variance totale pour l'équation (IV-11) appliquée avec tous les effets aléatoires est donnée par l'équation (IV-12).

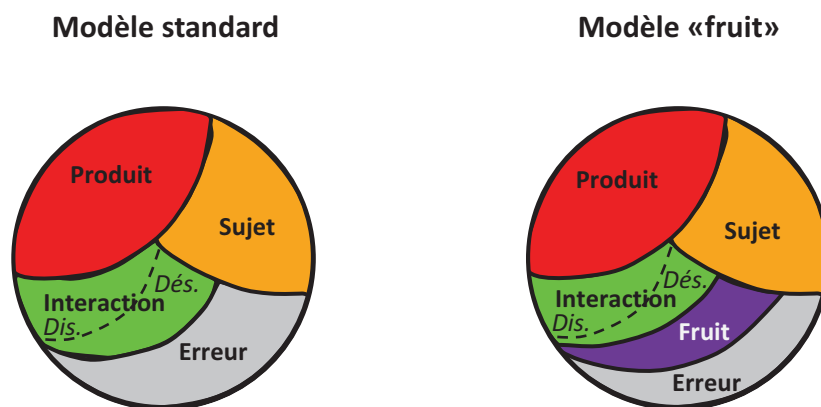
$$Var(X) = \sigma_{\text{variété}}^2 + \sigma_{\text{sujet}}^2 + \sigma_{\text{sujet*variété}}^2 + \sigma_{\text{fruit}}^2 + \sigma_e^2 \quad (\text{IV-12})$$

1.1.3 Comparaison des résultats des deux modèles

La comparaison des modèles a d'abord porté sur les conclusions concernant la discrimination ($F_{\text{variété}}$), le désaccord total entre les sujets ($F_{\text{sujet*variété}}$) et la répétabilité (σ^2).

De plus la décomposition de la variance a été observée pour les deux modèles. Pour cela, les modèles (IV-10) et (IV-11) ont été appliqués en considérant tous les effets comme aléatoires (même l'effet variété) afin de calculer les contributions à la variance de chaque facteur. Les variances totales pour les équations (IV-10) et (IV-11) appliquées avec tous les effets aléatoires sont données par les équations (IV-13) et (IV-12). La Figure 15 illustre la décomposition de la variance explicitée dans ces équations.

$$Var(X) = \sigma_{\text{variété}}^2 + \sigma_{\text{sujet}}^2 + \sigma_{\text{sujet*variété}}^2 + \sigma_e^2 \quad (\text{IV-13})$$



Dis. = dispersion et Dés. = désaccord pur

Figure 15. Décomposition de la variance selon le modèle standard et le modèle « fruit » (d'après Pineau (2006))

1.2 Résultats et Discussion

1.2.1 Conséquence de l'ajout du facteur « fruit » dans le cas d'un jeu de données réelles

Les données de la mesure de discrimination inter-variété ont ici été réutilisées. Ce jeu de données correspond à trois variétés évaluées quatre fois par 19 sujets selon sept descripteurs. Chaque fruit a été dégusté par un à quatre sujet(s). Les résultats de l'analyse des données de cette expérimentation avec le modèle standard sont présentés au Chapitre I.1.3.

Le Tableau 15 contient les résultats d'analyse de la variance du modèle « fruit » (IV-11). L'effet du facteur fruit est significatif au seuil de 0,01% pour les descripteurs *croquant*, *fermeté*, *crissant*, *jutosité* et *fondant*, au seuil de 5% pour le descripteur *acidité* et non significatif pour le descripteur *sucré* au seuil de 5%. L'importance de l'intégration du facteur fruit dans l'analyse de la variance est ici confirmée.

L'observation des contributions de chaque facteur à la variance (Tableau 16 illustré en Figure 16) indique que la contribution du facteur variété est prépondérante pour les descripteurs *croquant* (35%), *fermeté* (45%), *crissant* (33%) et *fondant* (35%). L'erreur est également importante. Elle est même prépondérante pour les descripteurs *jutosité* (37%), *acidité* (39%) et *sucré* (47%). Mis à part les contributions de l'erreur et du facteur variété, le fruit présente la plus forte contribution à la variance pour les descripteurs *croquant* (24%), *jutosité* (20%) et *fondant* (18%). La contribution du facteur fruit est supérieure à celle de l'interaction pour les descripteurs *fermeté* (15%), *crissant* (11%), *acidité* (8%) et *sucré* (7%). Aussi, quel que soit le descripteur, la contribution du fruit à la variance est supérieure à la contribution du désaccord total. Dans cette étude, le désaccord général entre les sujets a donc une importance moindre que la variabilité biologique des fruits. Le facteur sujet présente en général une contribution conséquente. Pour les descripteurs *acidité* et *sucré*, le facteur sujet est même le facteur qui présente la plus forte contribution (40% et 44% pour *acidité* et *sucré* respectivement). Il est vrai que tout au long des séances de panel, les sujets ont toujours montré plus de difficulté à s'accorder sur le niveau de notation pour ces descripteurs.

La comparaison des indices de discrimination, désaccord et répétabilité pour les deux modèles montre qu'un effet fruit fortement significatif ($p\text{-value} < 1,00E-03$) implique des variations pour ces indices.

Pour tous les descripteurs, excepté *sucré*, la discrimination ($F_{\text{variété}}$) baisse fortement lorsque nous passons du modèle standard au modèle « fruit » (Figure 17). Ce résultat s'interprète comme une diminution de la capacité de discrimination du panel. Cette diminution est expliquée par l'augmentation du bruit dans le ratio signal/bruit. En effet, avec le modèle standard, l'effet variété est testé contre le terme d'interaction sujet*variété alors qu'avec le modèle « fruit » il est testé contre une combinaison de l'interaction et de l'effet fruit. Un effet fruit extrêmement fort pourrait impliquer un changement de conclusion quant à la différence entre les produits. Dans le cas présenté ici, la discrimination est très forte et l'impact de l'effet fruit ne change pas les conclusions générales.

L'indice de répétabilité (σ^2) voit sa valeur diminuer pour tous les descripteurs (Figure 18). La chute observée pour le descripteur *sucré* est cependant moins forte. Cet effet est interprété comme l'augmentation de la répétabilité du panel. Cette meilleure répétabilité observée s'explique par la présence de l'effet fruit. En effet avec le modèle standard (IV-II-1-1), l'effet du facteur fruit semble reporté dans l'erreur puisque le passage du modèle standard au modèle « fruit » implique une forte chute de la part de variance non-expliquée (Tableau 10 et Tableau 16) (e.g. pour *croquant* : 53% > 24%, et pour *acidité* 48% > 39% avec le modèle standard et le modèle « fruit » respectivement).

L'interaction sujet*variété correspondant au désaccord total subit quant à lui une légère augmentation (excepté pour le descripteur *acidité*) (Figure 19). Cette augmentation s'explique par la diminution du terme d'erreur résiduelle, correspondant au bruit du ratio signal/bruit.

L'effet sujet augmente (Figure 20). Ce phénomène s'explique par la diminution du terme d'erreur qui constitue le bruit dans le ratio signal/bruit permettant de tester cet effet sujet.

Les résultats montrent que l'ajout de l'effet fruit au modèle standard implique des changements pour l'ensemble des paramètres statistiques observés ($F_{\text{variété}}$, $\text{Chi}^2_{\text{sujet}}$, $\text{Chi}^2_{\text{sujet*variété}}$ et σ^2). De plus, ils confirment l'hypothèse que la variabilité biologique des fruits est une composante importante de la variance de la réponse sensorielle. Cependant, dans le cas de ce jeu de données réelles, les conclusions sur les différences entre les produits (variétés) sont inchangées bien que la discrimination ait diminué. Ceci est probablement expliqué par une différence très marquée entre les variétés, qui n'est pas remise en cause par la variabilité intra-lot des fruits.

Tableau 15. Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs

Descripteurs	Variété		Sujet		Sujet*Variété		Fruit		Erreur σ^2
	$F_{2,67}$	p-value	Chi ²	p-value	Chi ²	p-value	Chi ²	p-value	
Croquant	27,30	2,04E-09***	39,62	3,09E-10***	2,32	1,28E-01	38,08	6,79E-10***	1,11
Fermeté	51,13	2,91E-14***	66,73	3,12E-16***	0,84	3,59E-01	32,32	1,31E-08***	1,22
Crissant	42,26	2,49E-12***	81,22	2,01E-19***	3,69	5,48E-02	12,95	3,20E-04***	1,15
Jutosité	14,02	1,07E-05***	9,29	2,30E-03***	4,01	4,52E-02*	21,42	3,69E-06***	1,43
Fondant	32,60	1,29E-10***	48,75	2,91E-12***	2,13	1,45E-01	25,34	4,81E-07***	1,57
Acidité	16,65	1,49E-06***	97,83	4,55E-23***	0,00	9,98E-01	5,23	2,22E-02*	1,57
Sucré	3,26	4,04E-02*	87,24	9,61E-21***	0,00	1,00E+00	2,97	8,46E-02	1,42

Niveau de significativité : 0,05* ; 0,01**, 0,001***

Tableau 16. Composantes de la variance et, entre parenthèses, leur pourcentage de la variation totale pour les sept descripteurs dans le cas du modèle « fruit » (IV-11) aléatoire

	Variété	Sujet	Interaction	Fruit	Erreur
Croquant	1,80 (35)	0,70 (13)	0,19 (04)	1,24 (24)	1,24 (24)
Fermeté	3,37 (45)	1,35 (18)	0,13 (02)	1,14 (15)	1,48 (20)
Crissant	1,82 (33)	1,47 (27)	0,26 (05)	0,60 (11)	1,32 (24)
Jutosité	1,22 (22)	0,75 (14)	0,40 (07)	1,09 (20)	2,04 (37)
Fondant	3,17 (35)	1,51 (16)	0,34 (04)	1,65 (18)	2,46 (27)
Acidité	0,84 (13)	2,57 (40)	0,00 (00)	0,50 (08)	2,46 (39)
Sucré	0,06 (01)	1,87 (44)	0,02 (01)	0,28 (07)	2,01 (47)

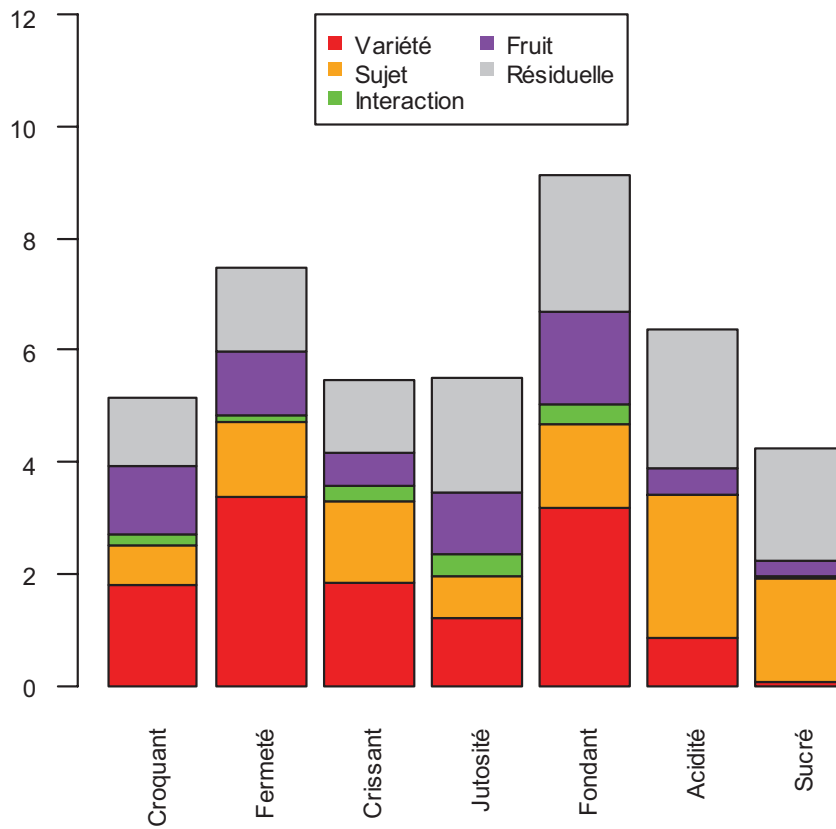


Figure 16. Estimations des composantes de la variance pour le modèle « fruit » (IV-11) et les sept descripteurs

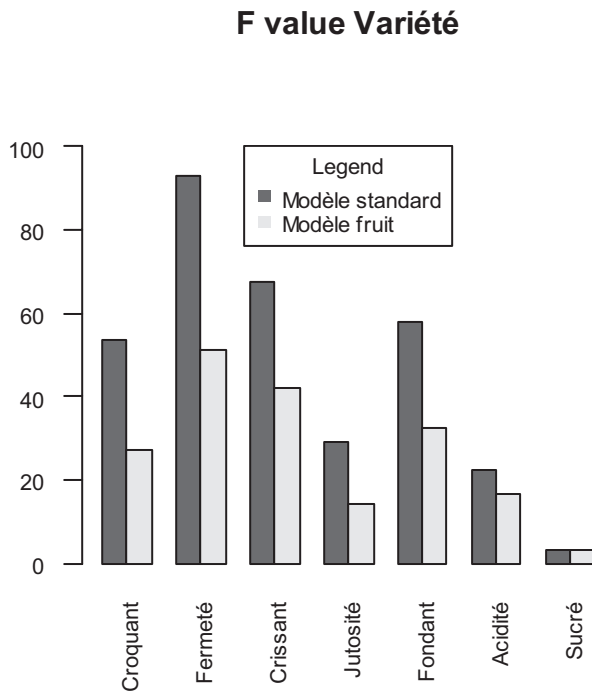


Figure 17. Passage du modèle standard au modèle « fruit » - discrimination

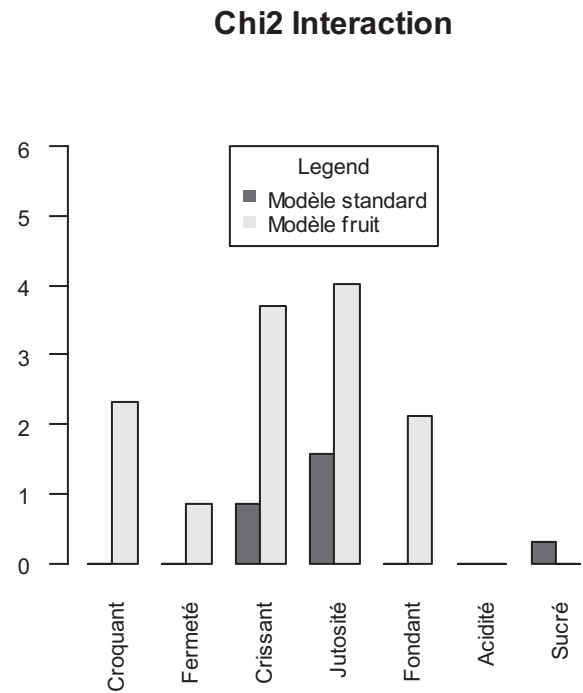


Figure 19. Passage du modèle standard au modèle « fruit » - désaccord total

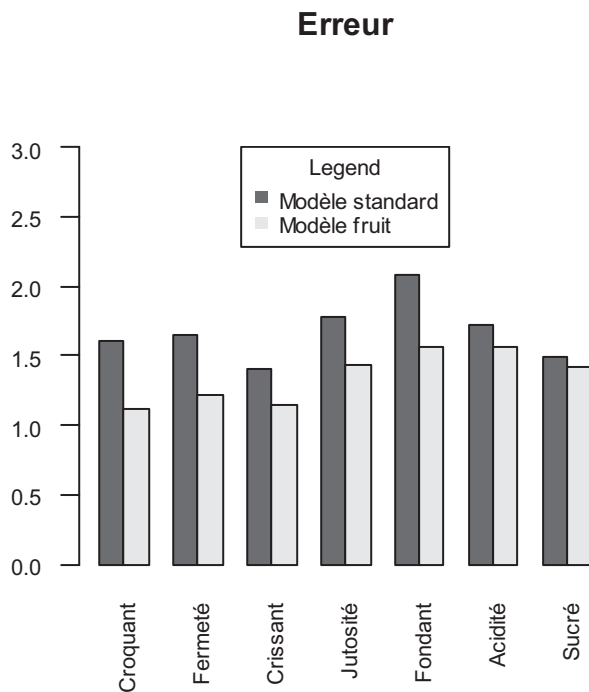


Figure 18. Passage du modèle standard au modèle « fruit » - répétabilité

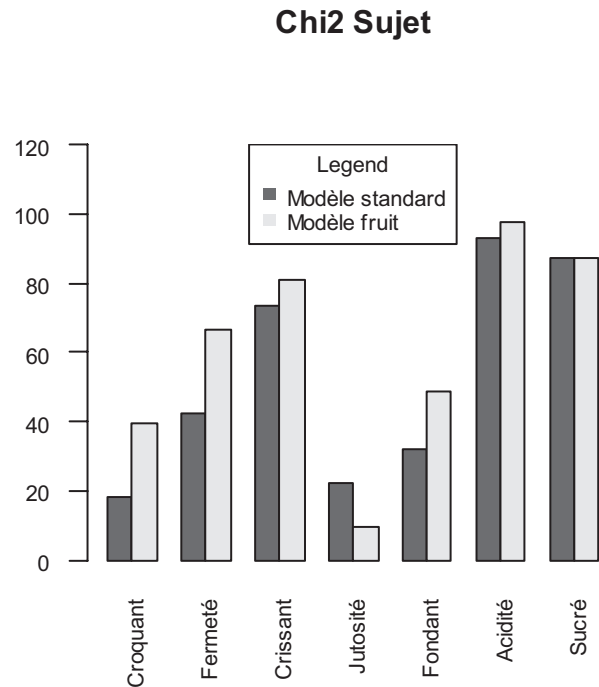


Figure 20. Passage du modèle standard au modèle « fruit » - effet niveau

1.2.2 Impact d'une variabilité intra-lot très élevée (données simulées)

Pour tester l'impact d'une forte variabilité intra-lot sur la différenciation de produits proches, des données ont été simulées.

Des données pour deux variétés, un panel de 20 sujets et deux répétitions pour chaque variété ont été simulées (Annexe 1). Deux scénarii de variabilité intra-lot ont été envisagés : variabilité intra-lot faible avec un écart-type égale à 0,8 (scénario 1) et variabilité intra-lot forte avec un écart-type égale à 2,5 (scénario 2). La loi Normale a été utilisée pour simuler les données avec une variété A de moyenne 4,5 et d'écart type 0,8 ou 2,5 et une variété B de moyenne 6,0 et d'écart-type 0,8 ou 2,5. Pour cette simulation, nous avons considéré que chaque fruit avait été dégusté par 2 sujets, il y a donc un total de 20 fruits par variété. Les données générées ne comportent pas d'interaction sujet*variété. Les distributions des deux variétés pour les deux scénarii sont représentées en Figure 21.

Les résultats d'ANOVA pour les deux scénarii et pour les deux modèles – le modèle standard et le modèle « fruit » – sont comparés.

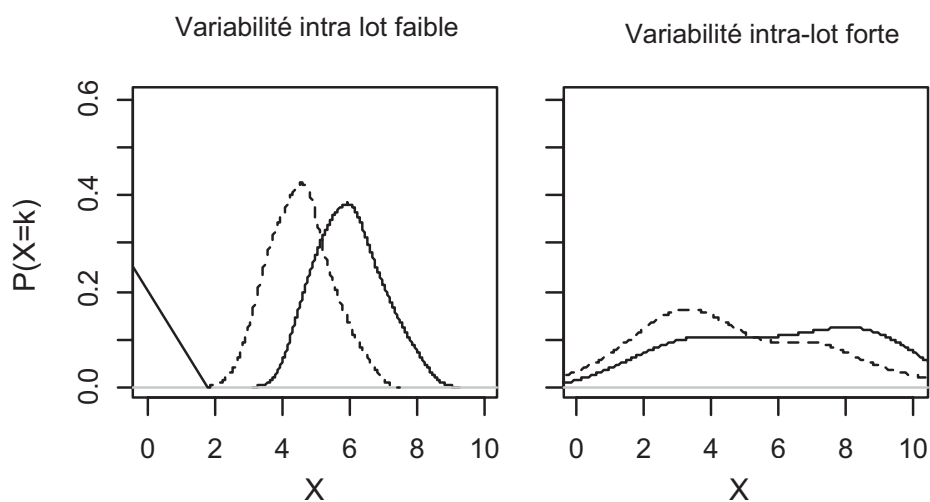


Figure 21. Distribution des données simulées pour des variabilités intra-lot faible (à gauche) et forte (à droite)

Les valeurs des statistiques F et χ^2 , les p-values associées et les contributions à la variance (cas d'un modèle aléatoire) sont reportées dans le Tableau 17. La Figure 22 et la Figure 23 illustrent les différences entre les résultats du modèle standard et du modèle « fruit », pour les scénarii 1 et 2 respectivement. L'analyse des deux types de données selon le modèle standard indique une discrimination ($F_{\text{variété}}$) des deux variétés quelle que soit la variabilité intra-lot. Cependant, la

discrimination est supérieure dans le cas de la variabilité intra-lot faible (scénario 1) ($67,81 > 6,83$). La valeur de l'erreur résiduelle est très élevée dans le cas de la variabilité intra-lot forte (scénario 2). Un tel résultat conduirait à la conclusion que le panel n'est pas répétable. Or, cela signifie aussi que la part de variabilité expliquée est faible. En effet, 75% de la variabilité reste non expliquée dans le cas de la variabilité intra-lot forte.

Lorsque le modèle « fruit » est appliqué, la discrimination diminue ($F_{\text{variété}}$) ($43,30 < 67,81$ et $3,51 < 6,83$ pour les variabilités faible et forte respectivement). Dans le cas de la variabilité intra-lot forte (scénario 2), la discrimination devient même non significative. Dans les deux cas, l'effet fruit est significatif. Concernant l'erreur, elle est réduite pour les deux scénarii. Pour le cas de la variabilité intra-lot forte (scénario 2), l'erreur est fortement réduite et la variabilité non-expliquée devient négligeable (2%).

Dans le cas d'une variabilité intra-lot forte, les conclusions tirées par l'expérimentateur changeraient en fonction du modèle d'analyse utilisé. Avec le modèle standard, l'expérimentateur conclurait à la différence des deux variétés, alors qu'avec le modèle « fruit », il conclurait à la non différence des deux variétés. Le choix du modèle prend ici toute son importance.

Tableau 17. Comparaison des résultats d'ANOVA dans deux situations de variabilité biologique

Modèle ANOVA		Scénario 1 – Variabilité intra-lot faible		Scénario 2 – Variabilité intra-lot forte	
		Modèle standard	Modèle « fruit »	Modèle standard	Modèle « fruit »
Variété	$F_{1,59}$ et $F_{1,35}$	67.81	43,30	6,83	3,51
	p-value	2,19E-11***	1,58E-07***	1,13E-02*	6,89E-01
	Contribution	55,40%	52,72%	12,39%	10,47%
Sujet	Chi^2	5,90	21,34	0,00	24,53
	p-value	1,51E-02*	3,84E-06***	1,00E+00	7,33E-07***
	Contribution	8,99%	17,42%	0,00%	4,89%
Sujet*Variété	Chi^2	0,67	0,00	0,88	0,38
	p-value	4,11E-01	1,00E+00	3,48E-01	5.37E-01
	Contribution	6,50%	0,00%	13,07%	0,39%
Fruit(Variété)	Chi^2	-	16,29	-	95,10
	p-value	-	5,41E-05***	-	1,81E-22***
	Contribution	-	19,99%	-	82,27%
Résiduelle	σ^2	0,73	0,44	2,34	0,39
	Contribution	29,12%	9,88%	74,54%	1,98%

Niveau de significativité : 0,05* ; 0,01** , 0,001***

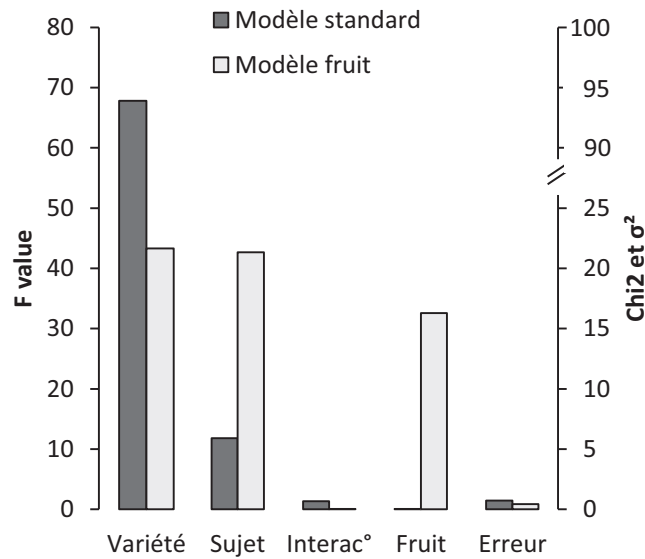


Figure 22. Statistiques de test pour les données simulées à variabilité intra-lot faible pour le modèle standard et le modèle fruit

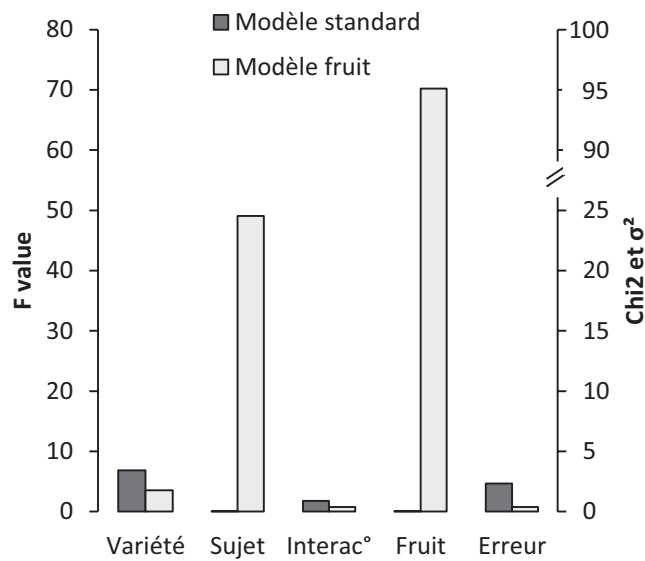


Figure 23. Statistiques de test pour les données simulées à variabilité intra-lot forte pour le modèle standard et le modèle fruit

1.3 Conclusion

L'application du modèle « fruit » aux données de mesure de discrimination inter-variétale a mis en évidence l'importance de la variabilité intra-lot. L'utilisation du modèle « fruit » tenant compte de la variabilité intra-lot a permis de réduire la part de variabilité non expliquée. Le modèle « fruit » entraîne donc une meilleure compréhension de la variabilité existante des données. En outre, l'addition de l'effet fruit génère de moins bons résultats pour la discrimination. Cette diminution de

la discrimination s'explique par l'augmentation de la partie bruit du ratio signal/bruit permettant de tester la significativité des différences entre les variétés.

Le modèle « fruit » devrait à notre avis se substituer au modèle standard pour toutes études de produits pouvant présenter une hétérogénéité. Son utilisation éviterait toutes conclusions trop optimistes quant à la différence entre les produits lorsque la variabilité intra-lot est conséquente.

Notons que l'introduction de cet effet fruit implique une conséquence au niveau de la méthodologie sensorielle appliquée pour le recueil des réponses. En effet, il est nécessaire que le nombre d'observations soit supérieur au nombre de fruits dégustés. Cette contrainte se traduit par la nécessité de faire déguster un même fruit par au moins deux sujets.

2. Étude de l'approche « assessor »

Lors de l'étude bibliographique, nous avons souligné l'interprétation erronée du terme d'interaction dans l'utilisation de l'analyse de la variance standard. Cette interaction est généralement interprétée comme le désaccord des sujets mais peut également être due aux différences de dispersion entre les sujets. Brockhoff et Skovgaard (1994) ont proposé un modèle d'analyse de la variance intégrant cet effet de notation, nommé « assessor model », mais celui-ci n'est pas exploité par les analystes sensoriels. L'intérêt de ce modèle a été démontré par Pineau (2006). La réflexion menée avec Per Brockhoff dans le cadre de ce travail de thèse sur la problématique d'adaptation du modèle d'analyse de la variance a conduit au souhait de coupler le modèle « assessor » et le modèle « fruit ». Cependant, pour des raisons techniques liées à la structure des données (déséquilibrés et bloc complexe), l'application directe du modèle n'est aujourd'hui pas possible. Une approche de ce modèle, nommée approche « assessor », a donc été réalisée par une transformation préalable des données.

Dans une première partie les différences de notation des sujets ont été étudiées. Ainsi, nous souhaitons identifier les variations décrites par Brockhoff (2003) et notamment les différences de niveau, les différences de dispersion et le désaccord pur dans le cadre d'une dégustation de trois variétés de pommes.

Suite à cette caractérisation des écarts de notation des sujets du panel, les données de discrimination inter-variétale (Chapitre I.1) ont été utilisées pour valider la pertinence de l'approche « assessor » du modèle « fruit ».

2.1 *Observation des différences de notation des sujets*

Plusieurs auteurs se sont préoccupés des différences de notations existant entre les sujets d'un panel. Afin de justifier l'utilisation de la décomposition de ces différences entre les sujets, les différences de notation des sujets, c'est-à-dire l'écart des notes de chaque sujet par rapport à la moyenne du panel, ont été caractérisées.

Notre hypothèse est que les différences entre les notes des sujets sont le résultat de la combinaison de différences de niveau, de différences de dispersion, et d'un certain désaccord malgré un entraînement dédié.

Une autre préoccupation concerne la constance des différences de notation entre les sujets : en fin de séance les différences de notation s'accroissent/s'atténuent-elles avec la fatigue (répétabilité) ? Les différences de notation sont-elles les mêmes d'une séance à une autre (reproductibilité) ?

2.1.1 Démarche expérimentale

Le Tableau 18 résume la démarche expérimentale. Les trois variétés de pommes – Ariane, Braeburn et Pink Lady® – ont été évaluées afin de comparer les différences de notation en fonction des variétés. Afin d'évaluer la répétabilité (étroitesse de l'accord entre les résultats successifs obtenus au cours d'une séance), l'évaluation de la variété Ariane a été répétée cinq fois. Les deux autres variétés ont été répétées deux fois seulement afin de limiter la fatigue des sujets (neuf évaluations au total). Ce protocole a été répété trois fois, la reproductibilité (étroitesse de l'accord entre les résultats successifs obtenus au cours de séances différentes) peut ainsi être évaluée. Pour cette caractérisation, 14 sujets étaient présents. Le panel a été divisé en deux sous-panels (A et B) de sept sujets. Ainsi pour chaque échantillon, deux pommes ont été coupées pour être partagées entre sept sujets. L'ordre de dégustation était le même pour tous les sujets.

Les évaluations ont été effectuées sur un seul morceau de pomme. C'est pourquoi, la liste des descripteurs était réduite (Partie III.2.1).

Tableau 18. Démarche expérimentale pour l'observation des écarts de notation

Ordre de présentation	Séance 1	Séance 2	Séance 3
1	Ariane	Ariane	Ariane
2	Ariane	Ariane	Ariane
3	Ariane	Ariane	Ariane
4	Ariane	Ariane	Ariane
5	Ariane	Ariane	Ariane
6	Braeburn	Braeburn	Braeburn
7	Braeburn	Braeburn	Braeburn
8	Pink Lady®	Pink Lady®	Pink Lady®
9	Pink Lady®	Pink Lady®	Pink Lady®

2.1.2 Sélection des fruits

Pour limiter l'intervention de la variabilité intra-lot, une sélection des fruits en fonction de leur fermeté acoustique a été réalisée. 180 à 200 fruits des trois variétés étudiées ont d'abord été évalués par sonométrie. Pour chaque variété, la moyenne de fermeté acoustique a été calculée. Les fruits sélectionnés pour l'évaluation sensorielle avaient une fermeté acoustique correspondant à la moyenne de la variété. Pour chaque variété, la fermeté acoustique moyenne est reportée dans le Tableau 19, ainsi que l'intervalle de fermeté acoustique des fruits dégustés par le panel.

Tableau 19. Écart de notation – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des fruits dégustés

Variété	Fermeté acoustique moyenne	Étendue de fermeté acoustique des fruits dégustés
Ariane	18,09 ± 2,72	[17,40 ; 18,70]
Braeburn	24,99 ± 2,83	[24,60 ; 25,20]
Pink Lady®	25,33 ± 1,43	[25,20 ; 25,50]

2.1.3 Modèles mixtes

Les notes données par chaque sujet ont été comparées à la moyenne des sept notes du fruit caractérisé par le sous-panel comprenant ce sujet et six autres sujets. Les différences à la moyenne du sous-panel ont ensuite été analysées au moyen d'ANOVAs. Dans un premier temps, la constance des différences de notation a été testée par rapport au facteur sujet et au facteur variété. Étant donné que les données sont des écarts à la moyenne, l'effet variété n'a pas de sens (il sera toujours nul). C'est l'interaction sujet*variété qui indiquera s'il y a un effet variété, à savoir si certains sujets ne surnotent ou ne sous-notent qu'une partie des variétés. Par ailleurs, l'objectif étant ici de

caractériser les sujets, l'effet du facteur sujet est fixe. L'effet du facteur variété pourrait alors être aléatoire mais notre hypothèse que les différences de notation diffèrent d'une variété à une autre implique de conserver cet effet en fixe (Pineau, 2006).

$$Y_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijr} \quad (\text{IV-14})$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \text{ indépendants entre eux}$$

Où μ correspond à l'effet moyen global, ν_i à l'effet fixe variété, α_j à l'effet fixe sujet, γ_{ij} à l'interaction sujet*variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Dans un second temps la répétabilité et la reproductibilité sont testées pour chaque variété individuellement. Les effets des facteurs séance (reproductibilité) et position (répétabilité) sont aléatoires.

$$Y_{ijps} = \mu + \alpha_j + \lambda_p + \Lambda_{jp} + \pi_s + \Pi_{js} + \varepsilon_{ijps} \quad (\text{IV-15})$$

$$\lambda_p \sim N(0, \sigma_{\text{position}}^2)$$

$$\Lambda_{jp} \sim N(0, \sigma_{\text{sujet*position}}^2)$$

$$\pi_s \sim N(0, \sigma_{\text{séance}}^2)$$

$$\Pi_{js} \sim N(0, \sigma_{\text{sujet*séance}}^2)$$

$$\varepsilon_{ijps} \sim N(0, \sigma_e^2) \quad \text{indépendants entre eux}$$

Où μ correspond à l'effet moyen global, α_j à l'effet fixe sujet, λ_p à l'effet aléatoire position, Λ_{jp} à l'interaction (aléatoire) sujet* position, π_s à l'effet aléatoire séance, Π_{js} à l'interaction (aléatoire) sujet*séance et ε_{ijps} à l'erreur résiduelle (aléatoire) du modèle.

2.1.4 Les différences de notation des sujets

Le Tableau 20 présente les résultats pour le modèle (IV-14). L'effet du facteur sujet est significatif pour tous les descripteurs. Ce résultat indique que les écarts de notation à la moyenne diffèrent d'un sujet à une autre. Certains sujets surnotent en moyenne alors que d'autres sous-notent en moyenne et les amplitudes de surnotation ou sous-notation varient. Ce résultat correspond à l'effet sujet classiquement observé en évaluation sensorielle.

L'interaction sujet*variété présente également un effet significatif pour tous les descripteurs. Certains sujets ne font pas le même écart de notation en fonction de la variété. La Figure 24 illustre les tendances de notation moyenne pour le descripteur *croquant* pour chaque sujet. À titre d'exemple, le sujet 465 montre une tendance à surnoter la variété Pink Lady® et à sous-noter légèrement les variétés Ariane et Braeburn par rapport à l'ensemble des sujets de son groupe. Le sujet 470 montre un écart de notation inverse : il sous-note la variété Pink Lady® et surnote les variétés Ariane et Braeburn. D'autres sujets ont un écart de notation constant quelle que soit la variété : par exemple le sujet 479 surnote et le sujet 469 sous-note. Enfin, parmi les sujets ayant un écart de notation constant, l'amplitude de cet effet peut varier en fonction de la variété. Par exemple, le sujet 467 surnote mais cet effet est plus important en moyenne pour la variété Braeburn. Pour un sujet donné, des écarts de notation à la moyenne de son groupe différents d'une variété à l'autre témoignent de l'effet dispersion.

Tableau 20. Effet du facteur sujet et de de l'interaction sujet*variété sur les tendances de notations pour les trois variétés de pomme et les sept descripteurs

Descripteurs	Facteur sujet		Interaction sujet*variété	
	F _{13 ; 336}	p-value	F _{28 ; 336}	p-value
Croquant	24,29	5,79E-41***	2,39	2,26E-04***
Ferme	37,70	3,52E-57***	2,57	6,37E-05***
Juteux	12,67	1,23E-22***	3,20	6,21E-07***
Fondant	24,93	7,38E-42***	4,87	1,70E-12***
Acide	21,36	1,00E-36***	2,72	2,21E-05***

Niveau de significativité : 0,001***

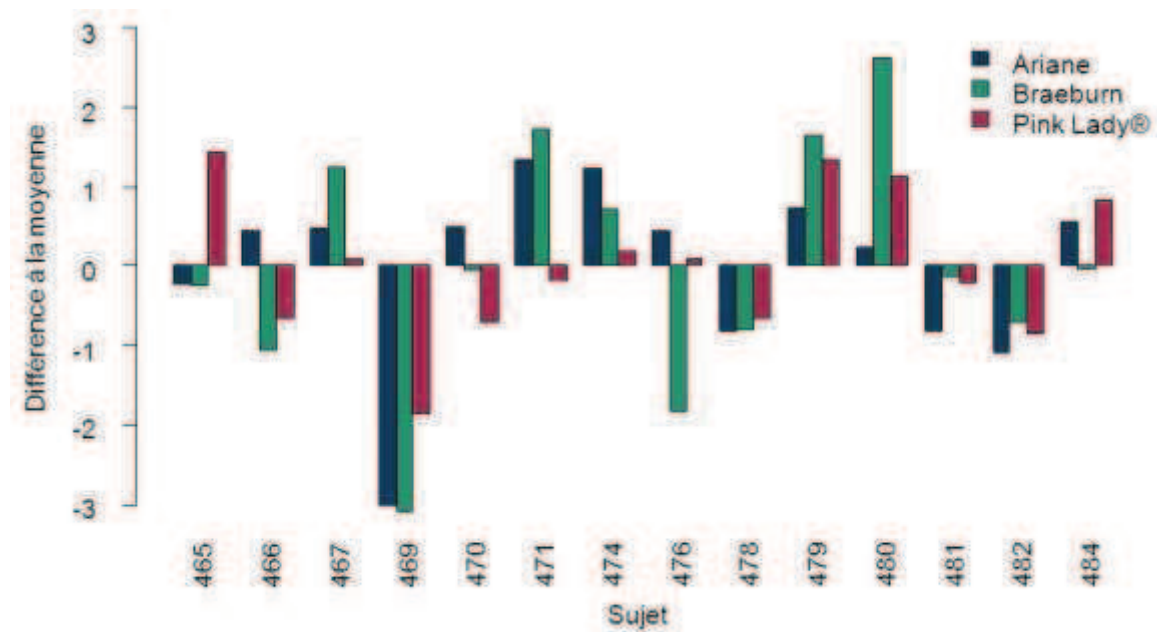


Figure 24. Différence à la moyenne les trois variétés de pomme et pour le descripteur *croquant*

La Figure 25 contient les évaluations moyennes par sujet par sous-panel et par séance pour le descripteur *croquant*. Cette représentation permet d'illustrer les différences entre les sujets et de visualiser l'effet de niveau (moyennes de notation différentes) et l'effet de dispersion (étendues de l'échelle de notation utilisées différentes). Les sujets 469 (sous-panel B) et 482 (sous-panel A) illustrent bien l'effet de niveau avec une sous-notation générale. Le sujet 469 illustre également l'effet dispersion (notes plus écartées). Les sujets 465 (sous-panel A) et 484 (sous-panel B) ne surnotent que les variétés de caractéristiques supérieures et montrent donc un effet niveau et un effet dispersion modérés. Dans le sous-panel B, le sujet 476 est en désaccord avec ses collègues pour deux des trois évaluations.

En outre, dans 64% des cas, le sujet j est en accord avec son groupe. Parmi les cas de désaccord (36%), 66% sont des inversions d'ordre 1 de produits perçus proche par le sujet et par son groupe. Il apparaît donc que dans le cas présent, les différences entre les sujets sont plus causées par des différences de niveau et de dispersion que par du désaccord. Ce résultat peut s'expliquer par l'utilisation d'un panel très entraîné dont les sujets sont en accord mais n'utilisent pas l'échelle de notation de manière identique (moyenne de notation différente et/ou dispersion différente). L'accord général des sujets a déjà été observé pour les mesures de discrimination inter et intra-variétale, de par la non significativité du terme d'interaction.

Concernant les quatre autres descripteurs, le même constat concernant la faible part de désaccord est réalisé. En effet, pour le descripteur *fermeté*, 69% de cas d'accord entre le sujet *j* et son groupe sont constatés et parmi les 31% de désaccord, 69% correspondent à des inversions d'ordre 1 de produits perçus proches par le sujet et par son groupe. Pour le descripteur *jutosité*, 67% de cas d'accord entre le sujet *j* et son groupe sont constatés et parmi les 33% de désaccord, 79% correspondent à des inversions d'ordre 1 de produits perçus proches par le sujet et par son groupe. Pour le descripteur *fondant*, 60% de cas d'accord entre le sujet *j* et son groupe sont constatés et parmi les 40% de désaccord, 88% correspondent à des inversions d'ordre 1 de produits perçus proches par le sujet et par son groupe. Pour le descripteur *acidité*, un plus fort désaccord est observé avec seulement 33% de cas d'accord entre le sujet *j* et son groupe. Toutefois, parmi les 67% de désaccord, 64% correspondent à des inversions d'ordre 1 de produits perçus proches par le sujet et par son groupe. Ainsi, quel que soit le descripteur, la plus grande partie des différences observées entre les sujets est due à l'effet de niveau et l'effet de dispersion. Comme pour le descripteur *croquant*, l'entraînement dédié du panel à la perception de faibles variations au sein de variétés peut expliquer ce résultat.

Nous observons maintenant la répétabilité et la reproductibilité des écarts de notation individuels à la moyenne du groupe pour chaque variété. L'effet de l'interaction sujet*position n'est jamais significatif, quel que soit le descripteur et quelle que soit la variété. Les écarts de notation sont donc répétables au cours d'une séance. D'une répétition à l'autre, un sujet qui présente un écart à la moyenne donné, le reproduit à la répétition suivante.

L'effet de l'interaction sujet*séance n'est significatif, au seuil de 5%, que pour la variété Pink Lady® et les descripteurs *jutosité* (p -value = 6,89E-03**) et *acidité* (p -value = 2,30E-02*). La reproductibilité des tendances de notation est observée, quelle que soit la variété, pour les descripteurs *croquant*, *fermeté* et *fondant*.

Ces résultats confirment la présence d'effet de niveau, de dispersion et de désaccord dans la réponse sensorielle (Brockhoff, 2003, Romano *et al.*, 2008). Le désaccord pur apparaît ici négligeable. Il serait donc judicieux de pouvoir distinguer dispersion et désaccord, aujourd'hui confondu dans le terme d'interaction (sujet*produit), afin de tester correctement la discrimination.

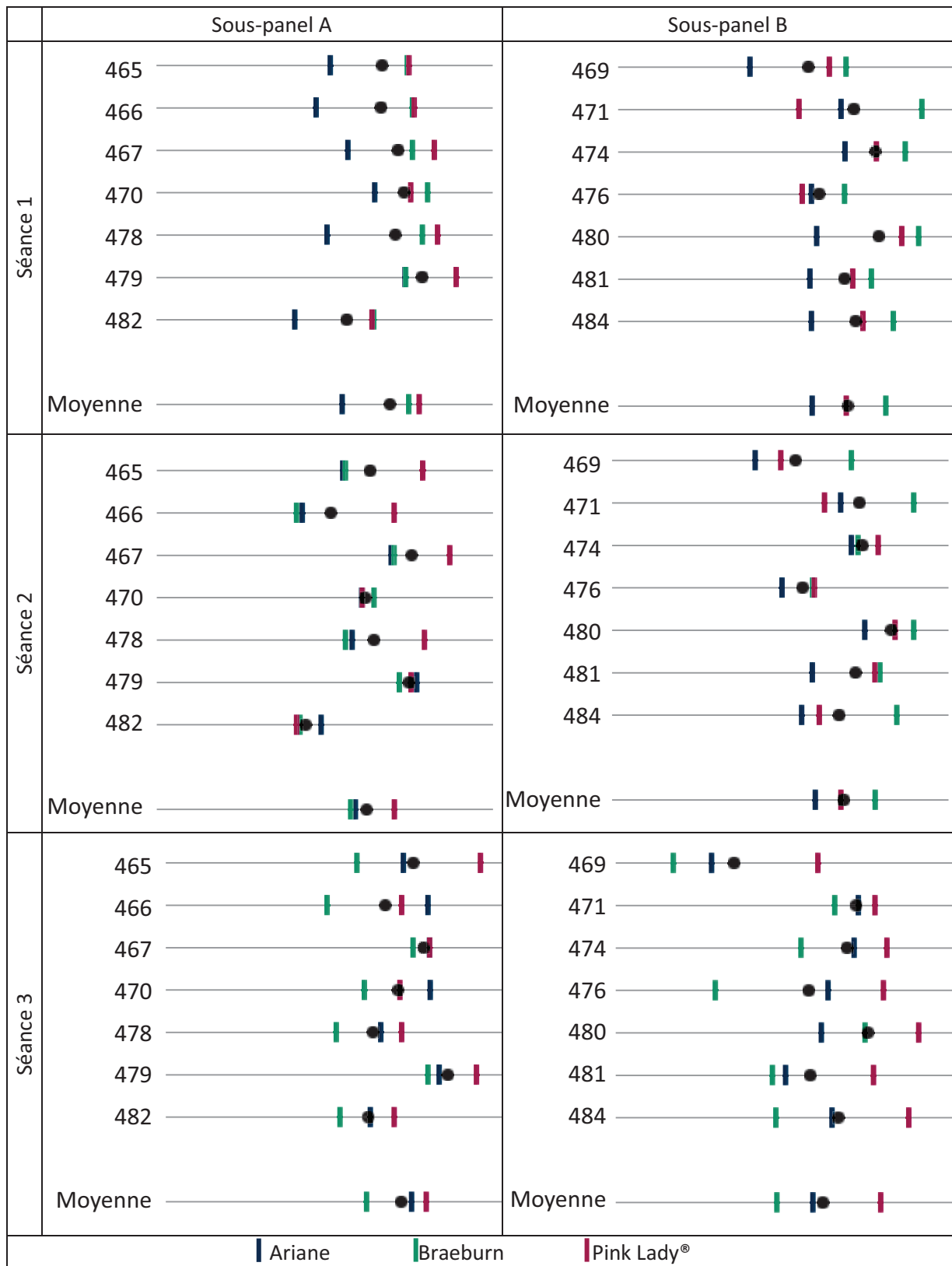


Figure 25. Évaluation moyenne individuelle des trois variétés pour les deux sous-panels et les trois répétitions pour le descripteur *croquant*

2.2 Approche « assessor »

Comme indiqué précédemment la distinction de la dispersion et du désaccord pur dans les modèles d'ANOVA doit permettre une meilleure estimation de la discrimination des produits étudiés. Les résultats issus du modèle « fruit » préalablement validé et ceux de l'approche « assessor » de ce même modèle sont maintenant comparés.

2.2.1 Les modèles

2.2.1.1 Le modèle « fruit »

Le modèle « fruit » est rappelé ci-après.

$$Y_{ijr} = \mu + \nu_i + \alpha_j + \gamma_{ij} + \varphi_{f(ijr)} + \varepsilon_{ijr} \quad (\text{IV-16})$$

$$\alpha_j \sim N(0, \sigma_{\text{ sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{ sujet*variété}}^2)$$

$$\varphi_{f(ijr)} \sim N(0, \sigma_{\text{ fruit(variété)}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \text{ indépendants entre eux}$$

Où μ correspond à l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), α_j à l'effet aléatoire sujet, γ_{ij} à l'interaction (aléatoire) sujet*variété, $\varphi_{f(ijr)}$ à l'effet aléatoire fruit imbriqué dans le facteur variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

2.2.1.2 L'approche « assessor »

Le modèle « assessor » a été proposé par Brockhoff et Skovgaard (1994) et consistait en un modèle comprenant uniquement les différences de dispersion en effet fixe. Une version étendue de ce modèle, utilisant les notes moyennes des produits en covariables, est proposée par Brockhoff *et al.* (2012). Ce modèle inclut un terme d'interaction sujet*produit représentant ici le désaccord pur dans la mesure où la dispersion est prise en compte par le facteur β_j :

$$X_{ijr} = \mu + \nu_i + \alpha_j + m_i \cdot \beta_j + \gamma_{ij} + \varepsilon_{ijr} \quad (\text{IV-17})$$

$$\alpha_j \sim N(0, \sigma_{\text{ sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{ sujet*produit}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \text{ indépendants entre eux}$$

Où μ correspond à l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), α_j à l'effet aléatoire sujet, m_i à la note moyenne du produit i , β_j à la dispersion du sujet j , γ_{ij} à l'interaction (aléatoire) sujet*variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Pour étudier conjointement la conséquence de la décomposition dispersion/désaccord et de l'effet du facteur fruit, idéalement, nous aimerions étendre le modèle avec l'effet aléatoire fruit :

$$X_{ijr} = \mu + \nu_i + \alpha_j + m_i \cdot \beta_j + \gamma_{ij} + \varphi_{f(ijr)} + \varepsilon_{ijr} \quad (\text{IV-18})$$

$$\alpha_j \sim N(0, \sigma_{\text{sujet}}^2)$$

$$\gamma_{ij} \sim N(0, \sigma_{\text{sujet*produit}}^2)$$

$$\varphi_{f(ijr)} \sim N(0, \sigma_{\text{fruit(variété)}}^2)$$

$$\varepsilon_{ijr} \sim N(0, \sigma_e^2) \text{ indépendants entre eux}$$

Où μ correspond à l'effet moyen global, ν_i correspond à l'effet fixe variété (lot), α_j à l'effet aléatoire sujet, m_i à la note moyenne du produit i , β_j à la dispersion du sujet j , γ_{ij} à l'interaction (aléatoire) sujet*variété, $\varphi_{f(ijr)}$ à l'effet aléatoire fruit imbriqué dans le facteur variété et ε_{ijr} à l'erreur résiduelle (aléatoire) du modèle.

Le traitement complet de ce modèle appliqué aux données non-équilibrées n'a pas pu être effectué. En effet, les outils statistiques permettant cette analyse ne sont pas encore développés, ni accessibles dans un logiciel. Notre approche a donc été la suivante : appliquer le modèle « fruit » (IV-16) à des données préalablement corrigées pour l'effet de dispersion.

La correction pour l'effet de dispersion est effectuée par une méthode additive similaire à celle proposée par Romano *et al.* (2008). L'effet de dispersion est estimé sur la base des moyennes centrées des produits :

$$X_{ijr} - m_i \cdot (\beta_j - \bar{\beta}) \quad (\text{IV-19})$$

Le terme $(\beta_j - \bar{\beta})$ correspond à la différence entre le coefficient de dispersion du sujet j et le coefficient de dispersion moyen. Il est positif si le sujet utilise une portion plus large de l'échelle. Le terme m_i est la moyenne du produit i (Pineau, 2006).

2.2.1.3 Comparaison des résultats des deux modèles

Dans un premier temps, pour chaque descripteur, l'évolution du terme d'interaction est observée, reflet du désaccord total dans le cas du modèle « fruit » et du désaccord pur dans le cas de l'approche « assessor ». Puis les indices de discrimination ($F_{\text{variété}}$) et de répétabilité (σ^2) issus des deux analyses sont comparés.

De plus, le modèle (IV-16) est appliqué en considérant tous les effets comme aléatoires (même l'effet variété), afin de calculer les contributions à la variance de chaque facteur, aux données brutes et aux données corrigées pour l'effet dispersion.

Les variances totales pour l'équation (IV-16) appliquée, avec tous les effets aléatoires, aux données brutes ou aux données corrigées sont données par les équations (IV-20) et (IV-21). La Figure 26 illustre la décomposition de la variance explicitée dans ces équations.

$$\text{Var}(X) = \sigma_{\text{variété}}^2 + \sigma_{\text{sujet}}^2 + \sigma_{\text{sujet*variété}}^2 + \sigma_{\text{fruit}}^2 + \sigma_e^2 \quad (\text{IV-20})$$

$$\text{Var}(X) = \sigma_{\text{variété}}^2 + \sigma_{\text{sujet}}^2 + \sigma_{\text{désaccord}}^2 + \sigma_{\text{fruit}}^2 + \sigma_e^2 \quad (\text{IV-21})$$

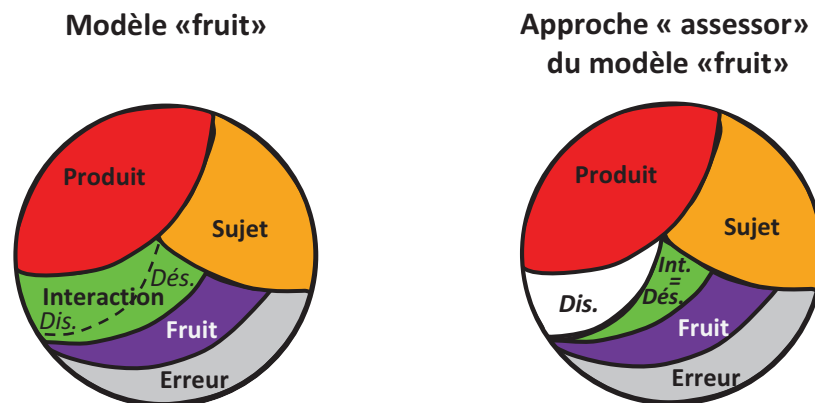


Figure 26. Décomposition de la variance selon le modèle « fruit » et l'approche « assessor » du modèle « fruit »

2.2.2 Résultats et Discussion

2.2.2.1 Conséquence de l'utilisation de l'approche « assessor » dans le cas d'un jeu de données réelles

Les données issues de la mesure de performance inter-variété ont été traitées avec le modèle standard et le modèle « fruit ». Ce jeu de données a donc été repris ici pour évaluer l'impact de l'approche « assessor » en comparaison avec le modèle « fruit ».

Le Tableau 21 contient les résultats de l'analyse des données de performance inter-variétale par l'approche « assessor » du modèle « fruit ». Les résultats indiquent un effet significatif pour le facteur variété et le facteur sujet pour tous les descripteurs. L'effet fruit est significatif sauf pour les descripteurs *acidité* et *sucré*. Quant au terme d'interaction, il est nul pour tous les descripteurs.

La comparaison des indices de discrimination, désaccord et répétabilité pour les deux modèles montre un impact majeur de l'approche « assessor » (Tableau 15 et Tableau 21).

Les termes d'interaction deviennent nuls lorsque l'approche « assessor » est utilisée, même pour le descripteur *jutosité* qui présentait une interaction significative (Figure 27). Cette annulation de l'interaction sujet*produit indique que le terme d'interaction présent dans les résultats précédent, alors interprété comme une combinaison du désaccord pur et de l'effet de dispersion, était entièrement dû à l'effet de dispersion. Ce résultat confirme à nouveau un très fort accord entre les sujet lié à un bon entraînement et que l'interaction observée n'est que le résultat de l'utilisation différente de l'échelle.

Pour tous les descripteurs, la valeur de la discrimination ($F_{\text{variété}}$) est supérieure avec l'approche « assessor » (Figure 28). Cette augmentation s'explique par la réduction du bruit du ratio signal/bruit suite à la diminution du terme d'interaction.

Le terme d'erreur est quant à lui peu changé mais tout de même plus faible pour l'approche « assessor » (Figure 29). La correction de la dispersion n'a un impact que sur l'interaction et pas la résiduelle.

L'importance de l'effet sujet augmente pour tous les descripteurs (Figure 30), probablement à cause de la diminution du terme d'interaction et de la légère diminution de l'erreur.

L'effet fruit (Figure 31) subit une légère augmentation pour les descripteurs *croquant*, *fermeté*, *crissant*, *jutosité* et *fondant*. Par contre pour les descripteurs *acidité* et *sucré*, l'effet fruit est diminué. De plus, la comparaison des composantes de la variance pour l'approche « assessor » du modèle « fruit » (Figure 32) avec celle du modèle « fruit » (Figure 16) montre une élimination du terme d'interaction. Dans l'approche « assessor » ce terme ne reflète plus que le désaccord pur.

Au final, il apparait donc que dans ce jeu de données avec un panel entraîné et très sensible, le terme d'interaction identifié dans le modèle « fruit » s'explique exclusivement par les différences de dispersion entre les sujets et non du désaccord entre les sujets. Les différences de résultats sur l'effet variété confirme la pertinence de l'intégration du modèle « assessor » dans l'analyse de la variance mais doivent être confirmée sur un jeu de données présentant plus de désaccord d'où l'utilisation d'un jeu de données simulées.

Tableau 21. Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs – performance inter-variété

Descripteurs	Variété		Sujet		Sujet*Variété		Fruit		Erreur
	F _{2,67}	p-value	Chi ²	p-value	Chi ²	p-value	Chi ²	p-value	σ ²
Croquant	30,17	4,37E-10***	44,31	2,81E-11***	0,00	9,99E-01	39,81	2,79E-10***	1,12
Fermeté	56,21	2,94E-15***	73,24	1,15E-17***	0,00	1,00E+00	34,33	4,66E-09***	1,20
Crissant	46,47	2,56E-13***	93,73	3,62E-22***	0,00	1,00E+00	17,00	3,75E-05***	1,13
Jutosité	18,52	3,91E-07***	41,70	1,06E-10***	0,00	1,00E+00	26,27	2,97E-07***	1,36
Fondant	34,10	7,13E-11***	60,81	6,30E-15***	0,00	1,00E+00	31,99	1,55E-08***	1,49
Acidité	24,82	2,18E-10***	103,81	2,23E-24***	0,00	9,96E-01	3,54	5,98E-02	1,51
Sucré	3,80	2,39E-02*	101,93	5,75E-24***	0,00	9,97E-01	0,40	5,29E-01	1,36

Niveau de significativité : 0,05* ; 0,01**, 0,001***

Chi2 Interaction

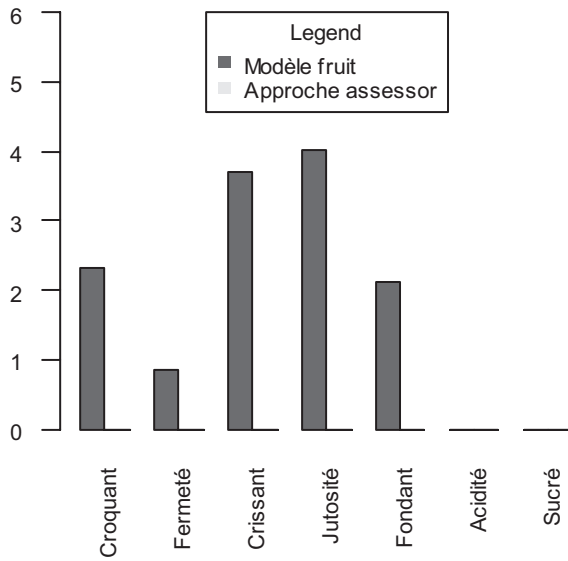


Figure 27. Passage du modèle « fruit » à l'approche « assessor » - désaccord total

Erreur

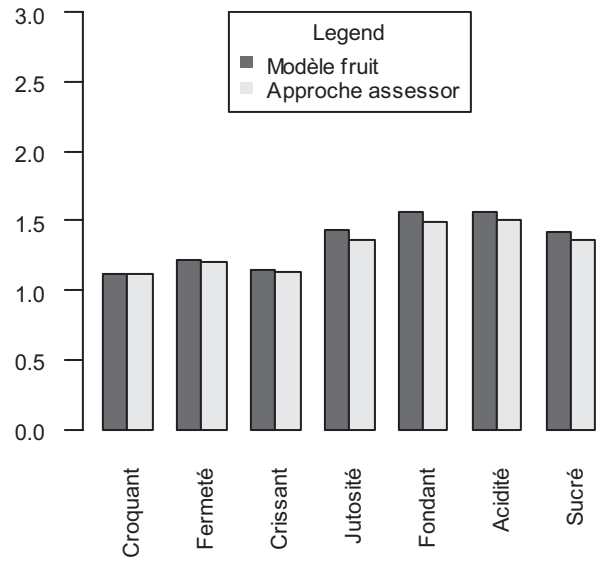


Figure 29. Passage du modèle « fruit » à l'approche « assessor » - répétabilité

F value Variété

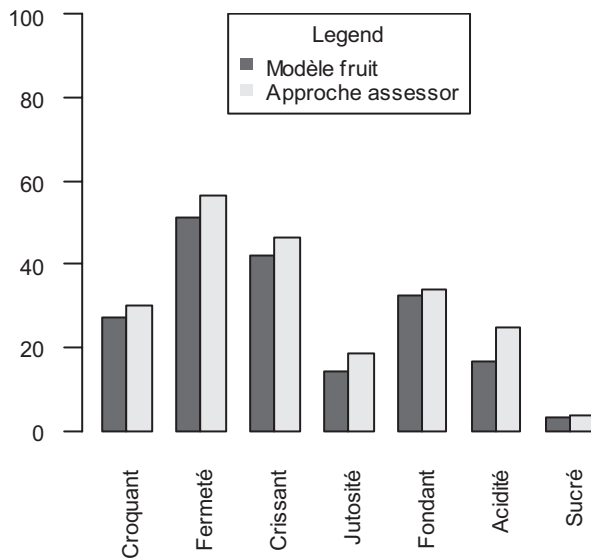


Figure 28. Passage du modèle « fruit » à l'approche « assessor » - discrimination

Chi2 Sujet

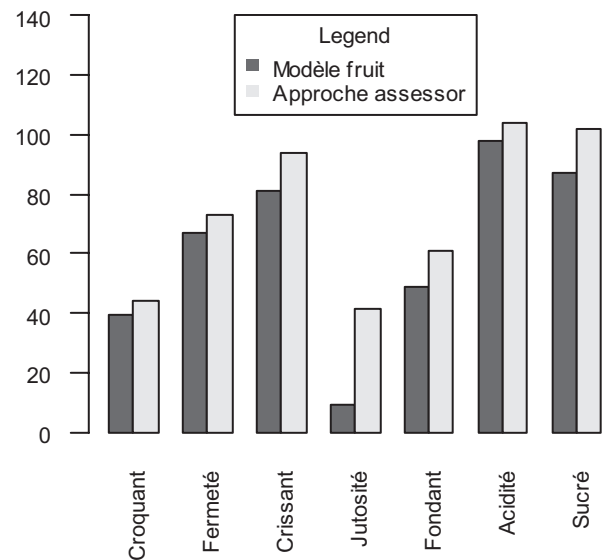


Figure 30. Passage du modèle « fruit » à l'approche « assessor » - effet de niveau

Chi2 Fruit

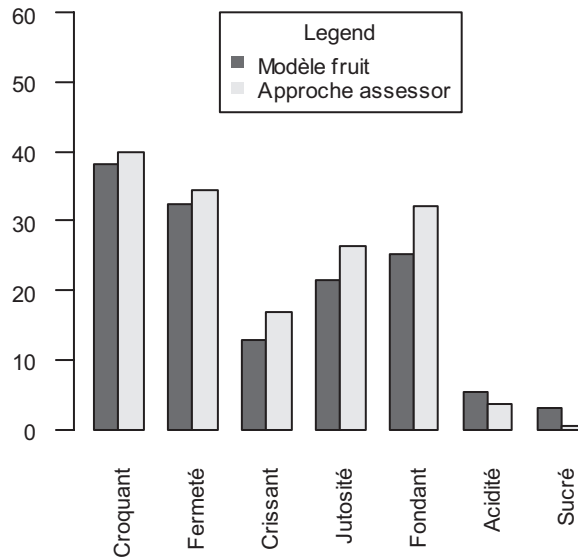


Figure 31. Passage du modèle « fruit » à l'approche « assessor » - effet fruit

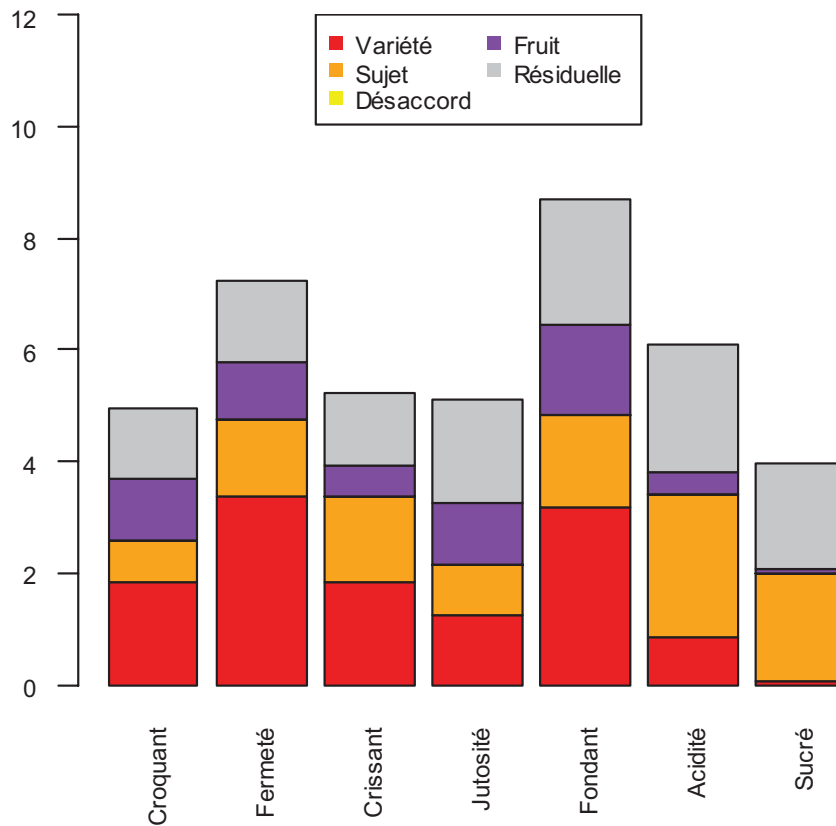


Figure 32. Estimations des composantes de la variance pour l'approche « assessor » du modèle « fruit » (IV-18) et les sept descripteurs

2.2.2.2 Impact de la présence de dispersion seule ou combinée à du désaccord

Afin d'observer les conséquences de l'utilisation du modèle incluant l'effet fruit et distinguant désaccord pur et dispersion, des données pour trois variétés, un panel de 20 sujets et deux répétitions pour chaque variété ont été simulées (Annexe 2). Deux scénarii d'interaction ont été envisagés. Dans les deux cas, un effet fruit a été généré. Dans le premier scénario, l'interaction est expliquée par des différences de dispersion entre les sujets uniquement et pas de désaccord (Figure 33). Dans le deuxième scénario, l'interaction est expliquée par du désaccord et des différences de dispersion (Figure 34). Pour ces simulations, nous avons considéré que chaque fruit avait été dégusté par deux sujets, il y a donc un total de 20 fruits par variété.

Les deux jeux de données générés sont analysés avec le modèle « fruit ». Parallèlement, l'approche « assessor » du modèle « fruit » est appliquée en réalisant d'abord la transformation décrite par l'équation (IV-19) pour éliminer l'effet de dispersion puis ces données corrigées sont analysées avec le modèle « fruit ». Les valeurs des statistiques F et Chi², les p-values associées et les contributions à la variance (cas d'un modèle aléatoire) sont reportées dans le Tableau 22. La Figure 35 et la Figure 36 illustrent les différences entre les résultats du modèle « fruit » et de l'approche « assessor » du modèle « fruit », pour les scénarii 1 et 2 respectivement.

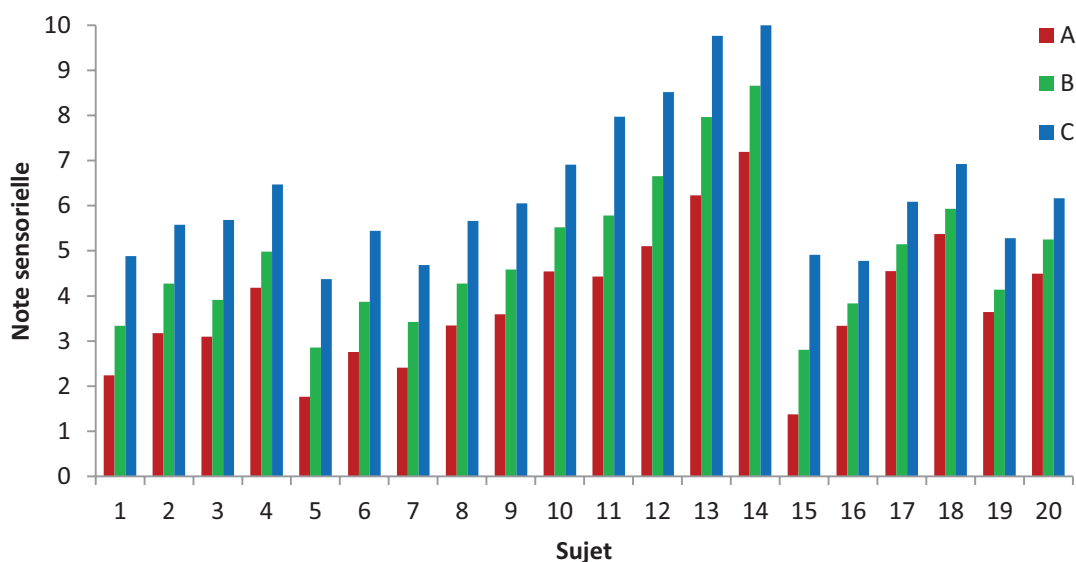


Figure 33. Scénario 1 avec dispersion et sans désaccord - Moyenne pour chaque sujet et chaque variété – les sujets 11 à 15 utilisent une plus large étendue de l'échelle ; les sujets 16 à 20 utilisent une plus faible étendue d'échelle

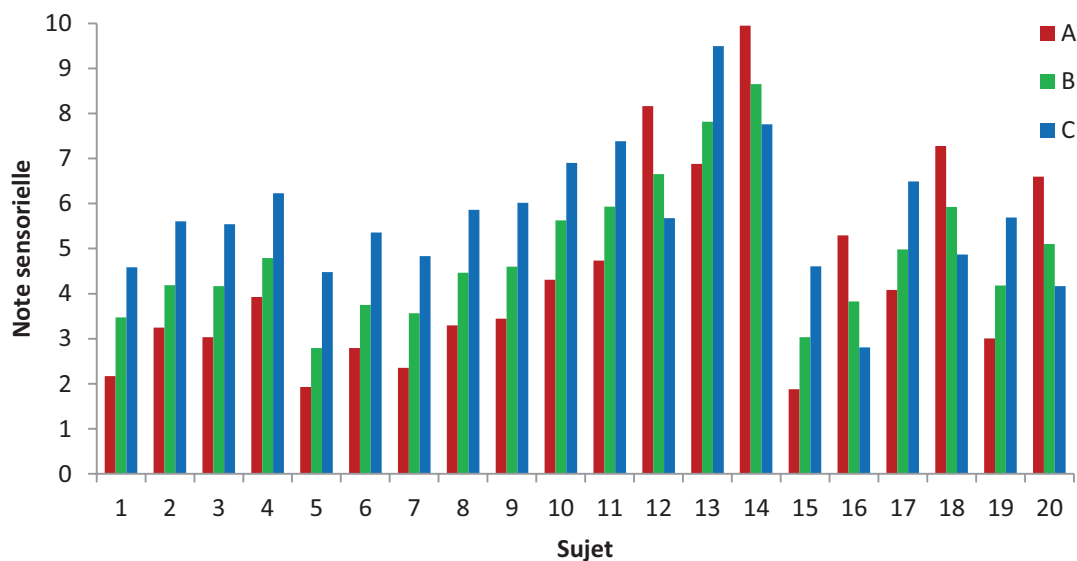


Figure 34. Scénario 2 avec désaccord et dispersion - Moyenne pour chaque sujet et chaque variété – les sujets 12, 14, 16, 18 et 20 présentent un classement des produits inverse au classement moyen (A = 4,4 < B = 4,9 < C =5,7)

Concernant les données présentant un effet de dispersion (scénario 1), l'analyse par le modèle « fruit » démontre une discrimination ($F_{\text{variété}}$) significative, une interaction sujet*variété significative, un effet fruit significatif et un effet sujet significatif. L'analyse par l'approche « assessor » entraîne l'annulation de l'interaction (qui devient non significative). Cela est dû à la transformation préalable des données qui a éliminé l'effet de dispersion, seul responsable de l'interaction observée par l'analyse avec le modèle « fruit ». La prise en compte de cet effet de dispersion par l'approche « assessor » implique une augmentation de la discrimination des variétés ($99,17 > 22,99$ respectivement pour l'approche « assessor » et le modèle « fruit »). Par ailleurs, la valeur de l'effet sujet augmente de manière très nette. Ceci est dû au fait qu'il est testé contre le terme d'interaction qui devient nul. Dans le cas d'une interaction nulle, il serait probablement préférable de tester l'effet variété contre l'effet fruit uniquement et l'effet sujet contre le terme d'erreur.

Concernant les données présentant du désaccord et de la dispersion (scénario 2), l'analyse par le modèle « fruit » indique une discrimination ($F_{\text{variété}}$) significative, une interaction significative, un effet fruit significatif et un effet sujet significatif. Cependant, l'interaction sujet*variété est supérieure à celle du scénario 1, à cause de la présence simultanée de désaccord pur et de dispersion. Cela a pour conséquence un effet variété plus faible puisque celui-ci est testé contre l'interaction. Le même constat est observé pour l'effet sujet. L'effet fruit, quant à lui, présente une valeur proche de celle observée pour le scénario 1. L'analyse par l'approche « assessor » entraîne

une diminution de la taille du terme d'interaction mais pas son annulation. En effet, dans ce cas et contrairement au scénario 1, l'interaction reflète la combinaison du désaccord pur et de l'effet de dispersion. La transformation appliquée aux données n'éliminant que la dispersion, une partie du terme d'interaction reste présente. Ce phénomène est illustré par les composantes de la variance en Figure 37. La prise en compte de l'effet de dispersion a pour conséquence l'augmentation de l'effet variété. Concernant, l'effet sujet, comme dans le scénario 1, sa valeur est augmentée par le biais de la diminution de l'interaction.

Tableau 22. Comparaison des résultats d'ANOVA pour l'effet de dispersion et le désaccord

Modèle ANOVA		Scénario 1 (dispersion)		Scénario 2 (désaccord et dispersion)	
		Modèle « fruit »	Approche « Assessor »	Modèle « fruit »	Approche « Assessor »
Variété	$F_{2;54}$ et $F_{2;54}$	70,55	99,17	5,09	22,99
	p-value	1,03E-15***	1,24E-17***	9,45E-03***	5,07E-08***
	Contribution	36,67	37,47	8,81	13,16
Sujet	χ^2	87,92	213,33	17,63	96,84
	p-value	6,80E-21***	2,57E-48***	2,69E-05***	7,51E-23***
	Contribution	52,48	54,68	47,88	47,59
Sujet*Variété	χ^2	13,43	0,00	109,42	21,01
	p-value	7,48E-04***	1,00	1,31E-25***	4,54E-06***
	Contribution	1,04	0,00	31,81	1,54
Fruit(Variété)	χ^2	77,19	109,34	77,50	76,05
	p-value	1,55E-18***	1,36E-25***	1,32E-18***	2,76E-18***
	Contribution	9,20	7,42	11,06	10,16
Résiduelle	σ^2	0.16	0.14	0.12	0.13
	Contribution	0.61	0.44	0.43	0.55

Niveau de significativité : 0,001***

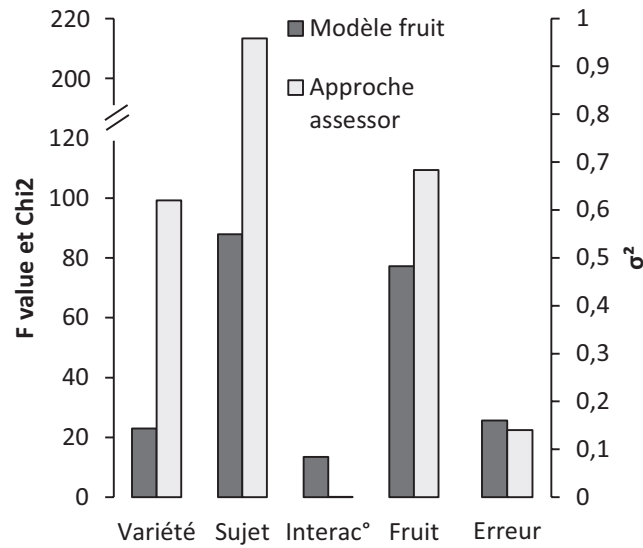


Figure 35. Statistiques de test pour les données simulées dans le cas du scénario 1 pour le modèle standard « fruit » et l'approche « assessor » du modèle « fruit ».

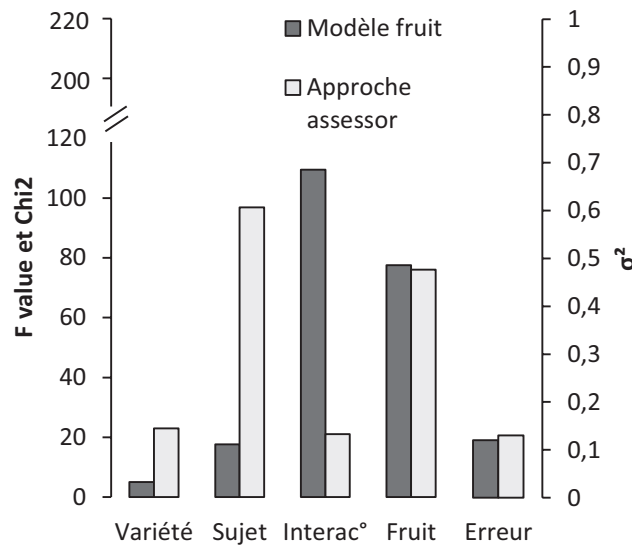


Figure 36. Statistiques de test pour les données simulées dans le cas du scénario 2 pour le modèle standard « fruit » et l'approche « assessor » du modèle « fruit ».

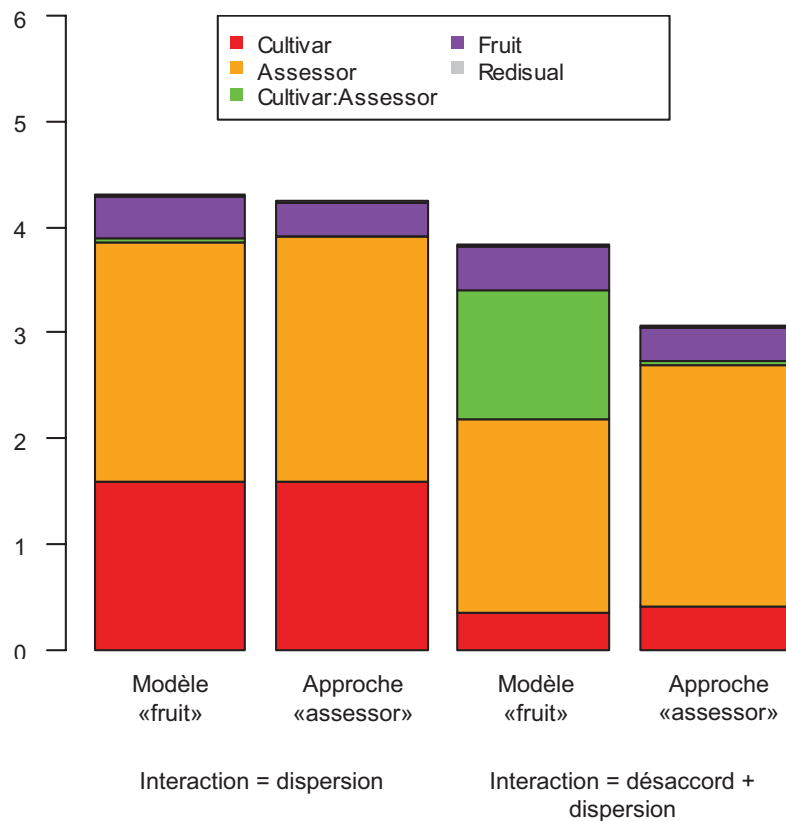


Figure 37. Estimation des composantes de la variance pour le modèle « fruit » et l'approche « assessor » du modèle « fruit » pour les données simulées scénario 1 (à gauche) et scénario 2 (à droite)

Pineau (2006) avait noté que la transformation donnée par l'équation (VI-19) implique une modification du classement des produits lorsque le sujet est en désaccord avec le panel. L'interprétation du paramètre β du modèle « assessor » comme l'indicateur de la dispersion individuelle des notes avait été remise en cause, le paramètre β pouvant être identique pour deux sujets présentant des dispersions différentes en fonction de leur accord avec le panel.

L'impact de la transformation (VI-19) a donc été vérifié dans le cadre des données simulées. Dans un premier temps les moyennes et les écart-types des notes et des notes transformées pour chaque sujet et chaque répétition sont calculées. Dans tous les cas, les moyennes sont inchangées. La transformation implique une réduction des écart-types de moitié (résultats non présentés).

Dans un second temps, afin de vérifier l'impact de la transformation sur le classement des produits, les coefficients de corrélations entre les notes et les notes transformées ont été calculés pour chaque sujet et chaque répétition (Tableau 23). Les coefficients de corrélation associés aux sujets présentant initialement un désaccord avec le panel (sujet 12, 14, 16, 18 et 20) sont négatifs. L'ordre des produits est donc modifié par la transformation. La comparaison des classements avant et après transformation confirme cette observation. Le classement initial $C > B > A$ pour les sujets en

désaccord (A > B > C du panel) devient B > A > C. Le traitement tend à aligner les notes des produits sur le classement du panel. L'observation faite par Pineau (2006) est ici confirmée. L'équivalence entre la transformation (VI-19) et le modèle tenant compte de l'hétérogénéité des dispersion (VI-17) avait été vérifiée par Pineau (2006). La pertinence du modèle est alors mise en question.

Tableau 23. Coefficients de corrélation de Pearson entre les notes et les notes transformées pour chaque sujet et chaque répétition (n=120)

Sujet	R(notes, note corrigées)		Sujet	R(notes, note corrigées)	
	Répétition 1	Répétition 2		Répétition 1	Répétition 2
1	0.99	0.98	11	0.98	1.00
2	0.99	1.00	12	-0.59	-0.76
3	0.98	1.00	13	1.00	1.00
4	1.00	1.00	14	-0.55	-0.51
5	1.00	1.00	15	1.00	0.99
6	1.00	1.00	16	-0.77	-0.64
7	0.99	0.99	17	1.00	1.00
8	0.99	0.99	18	-0.79	-0.76
9	0.99	0.99	19	0.99	0.99
10	0.99	0.98	20	-0.72	-0.60

2.3 Conclusion

L'observation des différences de notations entre les sujets du panel a montré une faible intervention du désaccord lors de l'évaluation des trois variétés de ce projet. Des différences de niveau et de dispersion ont été observées et les sujets sont globalement en accord.

L'approche « assessor » permet de distinguer l'effet de dispersion et le désaccord, confondus dans le terme d'interaction sujet*produit des modèles classiques. La soustraction de l'effet dispersion au terme d'interaction a pour conséquence des valeurs de discrimination supérieures. L'augmentation de la valeur de l'effet variété s'explique par la diminution de la valeur du terme d'interaction contre lequel il est testé. Dans le cas d'une interaction uniquement due à l'effet de dispersion, le terme d'interaction devient nul et les effets variété et sujet pourraient être exacerbés. Dans cette situation, l'utilisation de l'effet fruit pour tester l'effet variété et du terme d'erreur pour tester l'effet sujet serait plus adaptée.

L'approche « assessor » semble efficace dans la détection et la distinction du désaccord pur et de l'effet de dispersion. Cependant, le constat fait par Pineau (2006) de la modification du classement

des produits par la transformation (IV-19) dans le cas de la présence de désaccord est confirmé. La transformation semble alors excessive.

3. Conclusion générale

L'importance de la variabilité biologique a été mise en évidence par la forte contribution du facteur fruit à la variance totale des résultats de l'analyse avec le modèle « fruit ». L'utilisation de ce modèle a permis de réduire la part de variabilité non expliquée, qui avait été observée dans les résultats d'analyse par modèle standard. Le modèle « fruit » permet une meilleure compréhension de la variabilité existante des données et procure des résultats plus robustes. Le modèle « fruit » devrait à notre avis se substituer au modèle standard pour toutes études de produits pouvant présenter une hétérogénéité.

La pertinence de l'approche « assessor » a été démontrée. En effet, la distinction entre effet de dispersion et désaccord pur apporte une information supplémentaire sur les causes des différences entre les sujets et les performances de ceux-ci. L'absence de désaccord pur observée dans les données utilisées est alors mise en évidence dans les résultats. Néanmoins, cette approche nécessite une validation suite à l'observation d'un impact excessif de la transformation appliquée.

Chapitre III. Diminuer la variabilité : échantillonnage et présentation des échantillons

Suite à la mise en évidence de la capacité du panel à percevoir des différences entre les fruits d'une même variété, l'intégration de ce paramètre a été explorée au Chapitre II et a conduit à la validation du modèle « fruit » comprenant un facteur tenant compte de la variabilité intra-lot. Dans ce même chapitre une approche distinguant les différences de dispersion des sujets et le désaccord entre les sujets basée sur le modèle « assessor » a été évaluée. La mesure de la variabilité de la réponse sensorielle a été sensiblement améliorée pour permettre une meilleure compréhension des données.

Ce nouveau chapitre traite de la réduction de la variabilité. En effet, l'étude bibliographique a mis en avant l'utilisation de l'échantillonnage sur la base d'une mesure instrumentale ou d'une présentation spécifique des fruits (e.g. un fruit pour plusieurs sujets, cubes mélangés, purée, etc.) pour réduire la variabilité. Cependant, l'efficacité des méthodes proposées n'a pas été étudiée. Ce troisième chapitre est donc consacré à l'étude de l'impact de la réduction *a priori* de la variabilité des lots sur celle de la réponse sensorielle, par le biais de l'homogénéisation.

Deux stratégies d'homogénéisation ont été traitées : l'échantillonnage instrumental des fruits et le partage d'un même fruit pour plusieurs sujets.

La première partie de ce chapitre évoque l'évaluation de l'efficacité d'un échantillonnage instrumental des fruits. Deux mesures instrumentales ont été évaluées : la pénétrométrie (mesure destructive) et la sonométrie (mesure non destructive). Pour chacune de ces mesures, l'efficacité de l'homogénéisation *a priori* a été évaluée en comparant les résultats d'évaluations sensorielles de lots hétérogènes et homogènes.

La deuxième partie présente l'évaluation d'une homogénéisation basée sur le partage des fruits entre plusieurs sujets. Trois méthodes de présentation ont été examinées : un fruit pour un sujet, un fruit pour quatre sujets et un fruit pour huit sujets.

Enfin la troisième et dernière partie présente un bilan de l'impact de l'homogénéisation des lots pour l'évaluation sensorielle.

1. Étude de l'impact d'une homogénéisation instrumentale sur la variabilité de la réponse sensorielle

Depuis une dizaine d'années, la recherche dans le domaine de l'évaluation sensorielle appliquée aux produits végétaux n'a cessé de progresser et la question de l'homogénéisation des échantillons a été soulevée par certains auteurs qui ont utilisé des mesures instrumentales comme outils d'homogénéisation des fruits en amont de l'évaluation sensorielle (Harker *et al.*, 2002a, Harker *et al.*, 2008, Konopacka et Plochanski, 2004). L'accent a été mis sur l'homogénéisation des échantillons sur la base des propriétés mécaniques des fruits ou de leur teneur en sucres et en acides, mesurées par des outils destructifs (pénétrométrie, teneur en sucre soluble, acidité titrable) ou par des outils non destructifs (sonométrie, microdéformation, spectroscopie visible proche infrarouge). Cependant, dans aucune des études, les conséquences de l'homogénéisation sur la qualité des résultats sensoriels n'ont été analysées. L'efficacité d'une homogénéisation instrumentale *a priori* reste donc à prouver, même dans le cas d'une relation démontrée entre la mesure instrumentale utilisée et les descripteurs sensoriels. Afin d'évaluer l'efficacité de l'échantillonnage instrumental, deux types d'échantillonnage ont été appliqués pour obtenir un lot hétérogène et un lot homogène issus de chaque variété. Les évaluations sensorielles des lots hétérogènes et homogènes ont ensuite été comparées.

1.1 Méthode d'échantillonnage

Deux mesures ont été sélectionnées pour l'échantillonnage. L'une est la mesure de référence pour déterminer la fermeté des pommes : la pénétrométrie. Cette mesure a été effectuée avec une machine de traction universelle (MTS Universal, Synergie 200H). La seconde mesure sélectionnée, la sonométrie, présente l'avantage d'être non destructive et est déjà appliquée pour le tri en ligne des fruits et légumes. Cette mesure a été effectuée avec un dispositif AFS (AWETA, Nootdorp, The Netherlands).

Pour tester l'efficacité de l'homogénéisation par mesure instrumentale, des profils sensoriels ont été réalisés sur des lots déterminés de manière instrumentale comme étant hétérogènes et homogènes, pour chaque variété. Pour déterminer les lots, un grand nombre de fruits des trois variétés étudiées a d'abord été évalué par pénétrométrie ou sonométrie (respectivement 200 et 600 fruits). Pour chaque variété, les fruits ont été classés selon leur fermeté mécanique F_f ou leur fermeté acoustique S et répartis en 5 groupes d'étendue égale (A à E par ordre de fermeté mécanique / acoustique moyenne croissante). Les lots qui ont été dégustés ont été prélevés dans les lots initiaux en suivant deux méthodes d'échantillonnage : prélèvement hétérogène et prélèvement homogène (Figure 38). La méthode hétérogène consistait à prélever autant de fruits dans les trois groupes A, C et E, pour l'échantillonnage par pénétrométrie, ou les cinq groupes déterminés précédemment, pour l'échantillonnage par sonométrie. Par opposition, la méthode homogène consistait à prélever les fruits uniquement dans le groupe moyen (groupe C) du lot initial. Les échantillonnages étant centrés sur la moyenne, les lots hétérogènes et homogènes avaient la même caractéristique instrumentale moyenne.

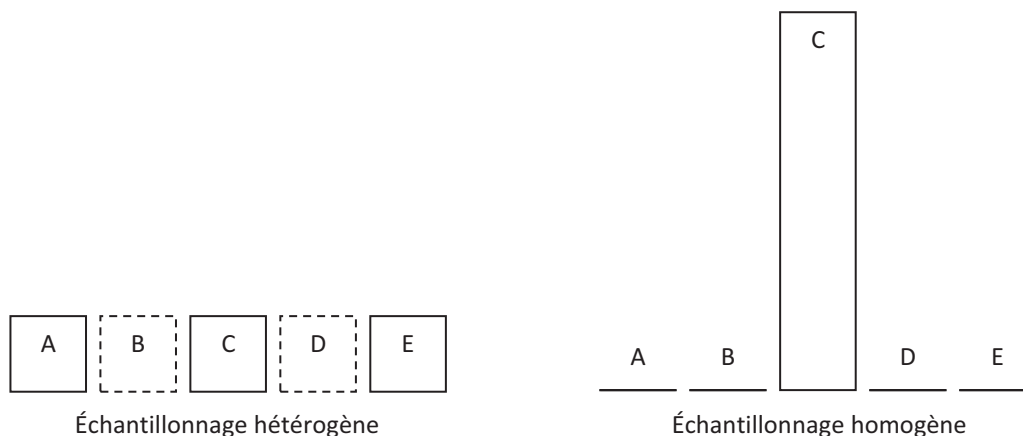


Figure 38. Méthodes d'échantillonnage hétérogène et homogène

Dans le cas de l'échantillonnage par pénétrométrie, lors de l'évaluation sensorielle, les sujets ont réalisé trois répétitions. Plusieurs sujets dégustaient le même fruit – deux à trois sujets par fruit. Un total de 18 fruits par variété a été évalué par profil. Pour chaque répétition, l'ordre des variétés a été basé sur un carré latin de Williams. 16 sujets ont pris part à cette mesure.

Dans le cas de l'échantillonnage par sonométrie, le même protocole a été mis en place à la seule différence que trois à quatre sujets partageaient le même fruit. Les 19 sujets ont pris part à cette mesure.

De plus, afin d'éviter un biais d'entraînement, le panel était divisé en deux sous-panels qui réalisaient les profils sur lots hétérogènes puis homogènes et inversement.

Pour chaque variété, la fermeté mécanique moyenne du lot initial, celle du lot hétérogène et celle du lot homogène sont reportées dans le Tableau 24. Les mêmes données concernant la mesure acoustique sont reportées dans le Tableau 25.

Tableau 24. Homogénéisation instrumentale – moyennes de fermeté mécanique (N) et écart-types des lots initiaux, hétérogènes et homogènes pour les trois variétés

Variété	Lot initial	Lot hétérogène	Lot homogène
Ariane	8,55 ± 1,10	8,62 ± 1,84	8,55 ± 0,09
Braeburn	9,18 ± 1,33	9,22 ± 2,30	9,18 ± 0,10
Pink Lady®	6,98 ± 0,96	7,00 ± 1,57	6,94 ± 0,07

Tableau 25. Homogénéisation instrumentale – moyennes de fermeté acoustique ($\text{Hz}^2\text{g}^{2/3}$) et écart-types des lots initiaux, hétérogènes et homogènes pour les trois variétés

Variété	Lot initial	Lot hétérogène	Lot homogène
Ariane	16,63 ± 4,98	18,93 ± 5,27	19,17 ± 0,30
Braeburn	22,85 ± 5,07	22,59 ± 5,71	22,01 ± 1,84
Pink Lady®	23,91 ± 3,14	23,05 ± 4,26	23,24 ± 0,21

1.2 Comparaison des échantillonnages

Le premier objectif de cette étude est de vérifier l'impact d'une réduction de la variabilité des lots dégustés sur la variabilité de la réponse sensorielle. La mesure de la variabilité de la réponse sensorielle a été effectuée par le calcul des écart-types pour chaque variété individuellement. Les écart-types obtenus pour les lots homogènes ont été comparés à ceux des lots hétérogènes avec un test de comparaison de variance de Fisher-Snedecor.

Par ailleurs, nous nous intéressons à l'impact de l'homogénéisation sur les résultats de l'ANOVA. En effet, nous supposons qu'un échantillonnage homogène aura pour conséquence une meilleure discrimination des variétés (lots). Les données issues des profils des deux types d'échantillonnages ont été analysées avec le modèle « fruit ». Les résultats ont ensuite été comparés.

1.3 Résultats & discussion

1.3.1 Échantillonnage destructif

Les écart-types de la réponse sensorielle pour chaque variété et chaque échantillonnage sont présentés dans le Tableau 26. Les écart-types des notes sensorielles pour les lots homogènes sont généralement plus faibles que ceux des lots hétérogènes, et notamment pour les variétés Ariane et Braeburn et les descripteurs de texture. La réduction des écarts-types varie selon les variétés et les descripteurs, avec une diminution minimale de 3% pour le descripteur *crissant* et la variété Ariane et une diminution maximale de 26% pour le descripteur *jutosité* et la variété Braeburn. Une différence variétale existe puisque l'homogénéisation ne semble pas avoir l'impact attendu pour la majorité des descripteurs de la variété Pink Lady®. L'analyse statistique des écarts-types montre néanmoins que cette réduction des écarts-types n'est pas significative excepté pour la variété Braeburn et le descripteur *jutosité* (p -value = 2,116E-02*), avec un écart-type plus faible dans le cas de l'évaluation du lot homogène. L'écart-type de la réponse sensorielle a donc tendance à être réduit par l'homogénéisation *a priori* du lot dégusté. Cette tendance mériterait d'être confirmée pour d'autres lots.

L'hypothèse de diminution de la variabilité de la réponse sensorielle par la réduction de la variabilité des lots dégustés par l'échantillonnage destructif est démontrée bien que les résultats ne soient que peu significatifs et différents en fonction des variétés. Le faible impact de l'homogénéisation par pénétrométrie sur la réduction de la variabilité de la réponse sensorielle peut s'expliquer par de faibles corrélations entre les descripteurs sensoriels et la fermeté mécanique. En effet, l'analyse des corrélations entre les descripteurs sensoriels et la fermeté mécanique (Tableau 27) indique que ces coefficients sont relativement faibles. Néanmoins les corrélations sont significatives au seuil de 5% pour la variété Ariane et tous les descripteurs excepté *sucré*, et la variété Pink Lady® et les descripteurs *croquant*, *fermeté*, *crissant* et *fondant* ; au seuil de 10% pour la variété Braeburn et les descripteurs *croquant* et *fermeté*. Les moindres corrélations observées pour la variété Braeburn confirment les résultats de Mehinagic (2004). De plus, les différences variétales observées concernant les relations entre descripteurs sensoriels et fermeté mécanique confirment la dépendance variétale des relations entre ces mesures observées dans la littérature (Mehinagic *et al.*, 2006, Molina-Delgado *et al.*, 2009a). Cependant ces différences de corrélations en fonction des variétés n'expliquent pas les différences variétales de l'impact de l'homogénéisation. Les coefficients de corrélations sont généralement plus faibles pour la variété Braeburn pour laquelle la réduction de la variabilité par l'homogénéisation est la plus forte.

Le Tableau 28 et le Tableau 29 contiennent les résultats d'analyse de la variance avec le modèle « fruit » pour le profil des lots hétérogènes et le profil des lots homogènes respectivement. Les changements des valeurs de discrimination en fonction de l'échantillonnage semblent aléatoires (augmentation pour quatre descripteurs et diminution pour trois descripteurs). Néanmoins, la valeur de discrimination ($F_{\text{variété}}$) pour le descripteur *jutosité* triple lorsqu'on passe du profil des lots hétérogènes au profil des lots homogènes ($4,93 < 15,64$). Pour ce même descripteur, l'effet fruit montre une diminution ($25,48 > 3,74$). Le terme d'interaction sujet*variété, qui n'est significatif que pour le profil des lots hétérogènes et les descripteurs *crissant* et *fondant*, ne subit pas de fort changement dans ce cas. Il est cependant non significatif quel que soit le descripteur pour le profil des lots homogènes. Dans les cas des descripteurs *crissant*, *jutosité* et *fondant*, la plus forte valeur de discrimination pour le profil des lots homogènes s'explique par la plus faible valeur de l'effet fruit. L'importante diminution de l'effet fruit suite à l'homogénéisation, dans le cas du descripteur *jutosité*, peut être mise en relation avec la forte diminution des écart-types pour les trois variétés. Au contraire, pour le descripteur *sucré*, pour lequel la performance était moins satisfaisante, la valeur de la discrimination ($F_{\text{variété}}$) diminue pour le profil des lots homogènes ($14,80 > 3,26$). Cet effet inattendu peut s'expliquer par l'augmentation de l'interaction sujet*variété dans le cas de l'échantillonnage homogène. Il est difficile d'émettre des hypothèses quant à cette évolution.

Tableau 26. Écart-types de la réponse sensorielle pour les trois variétés, sept descripteurs et deux échantillonnages par pénétrométrie, et évolution des écart-types

Variété	Ariane			Braeburn			Pink Lady®		
	Hétérogène.	Homogène.	Évolution.	Hétérogène.	Homogène.	Évolution.	Hétérogène.	Homogène.	Évolution.
Croquant	1.65	1.53	-7%	2.29	1.96	-14%	1.70	1.71	1%
Fermeté	1.81	1.59	-12%	2.04	1.85	-9%	1.72	1.66	-4%
Crissant	1.56	1.50	-3%	2.27	1.82	-20%	1.70	1.75	3%
Jutosité	2.03	1.80	-12%	2.43	1.80	-26% *	1.90	1.51	-20%
Fondant	2.13	1.99	-7%	2.82	2.62	-7%	2.12	2.30	8%
Acidité	1,95	1,98	1%	1,80	2,08	15%	2,14	1,95	-9%
Sucré	1,87	1,70	-9%	1,84	1,67	-9%	1,85	2,01	8%

Niveau de significativité : 0,05*

Tableau 27. Coefficient de corrélation de Pearson entre les descripteurs sensoriels et la fermeté mécanique (n=18)

	Croquant	Fermeté	Crissant	Jutosité	Fondant	Acidité	Sucré
Ariane	0,60*	0,71*	0,56*	0,53*	-0,63*	0,48*	0,22
Braeburn	0,43 ^(*)	0,41 ^(*)	0,17	-0,04	-0,24	0,19	0,20
Pink Lady®	0,53*	0,76*	0,57*	0,03	-0,57*	0,23	0,23

Niveau de significativité : 0,10^(*) ; 0,05*

Tableau 28. Échantillonnage hétérogène par pénétrométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs

Hétérogène	Variété		Sujet		Sujet*Variété		Fruit		Erreur σ^2
	$F_{2;76}$	p-value	Chi ²	p-value	Chi ²	p-value	Chi ²	p-value	
Croquant	18,49	2,00E-07***	27,98	1,23E-07***	0,00	1,00E+00	20,42	6,23E-06***	1.26
Fermeté	24,30	4,51E-09***	17,68	2,61E-05***	0,57	4,50E-01	14,07	1,76E-04***	1.30
Crissant	6,74	2,78E-03***	13,16	2,85E-04***	4,49	3,41E-02*	17,59	2,75E-05***	0.99
Jutosité	4,93	9,39E-03***	35,76	2,24E-09***	0,75	3,86E-01	25,48	4,46E-07***	1.23
Fondant	10,39	1,47E-04***	1,60	2,06E-01	15,95	6,52E-05***	19,05	1,27E-05***	1.81
Acidité	3,84	2,41E-02*	66,22	4,44E-16***	0,00	9,96E-01	0,45	5,03E-01	1.50
Sucré	14,80	1,69E-06***	77,86	1,11E-18***	0,00	9,98E-01	0,01	9,10E-01	1.25

Niveau de significativité : 0,05* ; 0,01**, 0,001***

Tableau 29. Échantillonnage homogène par pénétrométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs

Homogène	Variété		Sujet		Sujet*Variété		Fruit		Erreur σ^2
	$F_{2;76}$	p-value	Chi ²	p-value	σ^2	p-value	Chi ²	p-value	
Croquant	14,64	3,88E-06***	33,64	6,63E-09***	0,00	1,00E+00	20,47	6,06E-06***	1.35
Fermeté	24,84	4,81E-09***	23,59	1,19E-06***	0,22	6,39E-01	16,28	5,46E-05***	1.24
Crissant	8,16	6,13E-04***	28,25	1,06E-07***	0,79	3,74E-01	16,17	5,78E-05***	1.16
Jutosité	15,64	8,57E-07***	25,60	4,19E-07***	1,16	2,81E-01	3,74	5,30E-02	1.55
Fondant	15,96	1,68E-06***	23,04	1,59E-06***	0,32	5,74E-01	13,45	2,45E-04***	2.39
Acidité	1,09	3,39E-01	48,09	4,08E-12***	0,51	4,76E-01	2,09	1,48E-01	1.69
Sucré	3,26	4,17E-02*	52,80	3,69E-13***	1,95	1,63E-01	0,00	1,00E+00	1.58

Niveau de significativité : 0,05* ; 0,01**, 0,001***

1.3.2 Échantillonnage non destructif

Les écart-types de la réponse sensorielle pour chaque variété et chaque échantillonnage sont présentés dans le Tableau 30. Contrairement aux résultats précédents, l'homogénéisation par l'outil acoustique semble plus efficace pour la variété Pink Lady® que pour les variétés Ariane et Braeburn. En effet, une réduction des écart-types (bien que non significative au seuil de 10% excepté pour *crissant*) est observée pour la variété Pink lady® et tous les descripteurs. Les valeurs des réductions varient en fonction du descripteur, allant de 2% pour le descripteur *jutosité* à 21% pour le descripteur *crissant*.

L'hypothèse de diminution de la variabilité de la réponse sensorielle par la réduction de la variabilité des lots dégustés par l'échantillonnage non destructif est démontrée mais les résultats ne sont pas significatifs et sont différents en fonction des variétés. Comme dans le cas de l'homogénéisation par pénétrométrie, le faible impact de l'homogénéisation par sonométrie sur la réduction de la variabilité de la réponse sensorielle peut s'expliquer par de faibles corrélations entre les descripteurs sensoriels et la fermeté acoustique. Effectivement, les corrélations entre les descripteurs sensoriels et la fermeté acoustique présentées dans le Tableau 31 sont faibles et non significatifs sauf pour la variété Braeburn et les descripteurs *croquant* ($\alpha = 0,10$), *fermeté* ($\alpha = 0,05$) et *jutosité* ($\alpha = 0,05$) et la variété Pink Lady® et les descripteurs *croquant* ($\alpha = 0,10$) et *jutosité* ($\alpha = 0,10$). Ici encore, les corrélations dépendent des variétés confirmant les résultats de (Molina-Delgado *et al.*, 2009a). Néanmoins, la dépendance variétale des corrélations entre mesures sensorielles et mesures acoustiques ne permet pas d'expliquer totalement les différences d'impact de l'homogénéisation observées en fonction des variétés.

De plus, la capacité du panel à différencier des fruits de fermeté acoustique basse et haute a été démontrée précédemment pour certains descripteurs et certaines variétés (Chapitre 1.2). Une étendue moins large de la mesure acoustique pourrait expliquer la divergence des résultats. Pour vérifier cette hypothèse, pour chaque variété, l'étendue de la mesure de fermeté acoustique couvrant l'ensemble des trois niveaux de fermeté acoustique pour la mesure de discrimination intra-variétale a été comparée à l'étendue de la mesure de fermeté acoustique du lot hétérogène. L'étendue de la mesure de fermeté acoustique n'est pas présentée pour les lots homogènes puisqu'elle est quasi nulle. Les étendues présentées dans le Tableau 32 ne sont pas très différentes. L'hypothèse avancée est donc invalidée.

Tableau 30. Écart-types de la réponse sensorielle pour les trois variétés de pomme, sept descripteurs et deux échantillonnages par sonométrie, et évolution des écart-types

Variété	Ariane			Braeburn			Pink Lady®		
	Hétérogène	Homogène	Évolution	Hétérogène	Homogène	Évolution	Hétérogène	Homogène	Évolution.
Croquant	1,53	1,63	6%	1,91	1,98	4%	1,70	1,41	-17%
Fermeté	1,49	1,90	27% ^(*)	1,71	1,75	2%	1,63	1,42	-13%
Crissant	1,48	1,50	1%	2,00	2,02	1%	1,64	1,30	-21% ^(*)
Jutosité	1,72	1,79	4%	2,08	2,10	1%	1,38	1,35	-2%
Fondant	2,05	1,99	-3%	2,42	2,55	5%	1,94	1,78	-8%
Acidité	1,88	1,93	3%	2,00	2,09	5%	1,88	1,66	-12%
Sucré	1,60	1,60	0%	1,72	1,63	-5%	1,69	1,69	0%

Niveau de significativité : 0,10^(*) ; 0,05***Tableau 31. Coefficient de corrélation de Pearson entre les descripteurs sensoriels et les mesures instrumentales (n=18)**

	Croquant	Fermeté	Crissant	Jutosité	Fondant	Acidité	Sucré
Ariane	0,15	0,10	0,28	0,34	-0,26	0,27	0,00
Braeburn	0,44 ^(*)	0,48*	0,28	0,47*	-0,28	-0,11	0,17
Pink Lady®	0,42 ^(*)	0,27	0,13	0,44 ^(*)	-0,17	0,26	-0,06

Niveau de significativité : 0,10^(*) ; 0,05*

Tableau 32. Étendues de fermeté acoustique pour les expérimentations de discrimination intra-variétale et d'homogénéisation par sonométrie.

	Discrimination intra-variétale - étendue couvrant les trois niveaux de fermeté acoustique	Homogénéisation par sonométrie - lot hétérogène
Ariane	[9,60 ; 29,50]	[11,80 ; 27,70]
Braeburn	[15,10 ; 35,40]	[11,80 ; 30,70]
Pink Lady®	[22,10 ; 32,00]	[17,00 ; 29,20]

Par ailleurs, une augmentation de l'écart-type est observée pour la variété Ariane et le descripteur *fermeté* ($1,49 < 1,90$) entre le lot hétérogène et le lot homogène. Ce résultat n'est pas en accord avec notre hypothèse initiale. Le graphique représentant les notes individuelles des sujets contre les notes moyennes par fruit pour la variété Ariane et le descripteur *fermeté* (Figure 39) montre une plus grande dispersion entre les fruits dans le cas de l'échantillonnage homogène. Cela remet à nouveau en cause la pertinence de la mesure acoustique pour l'homogénéisation des lots.

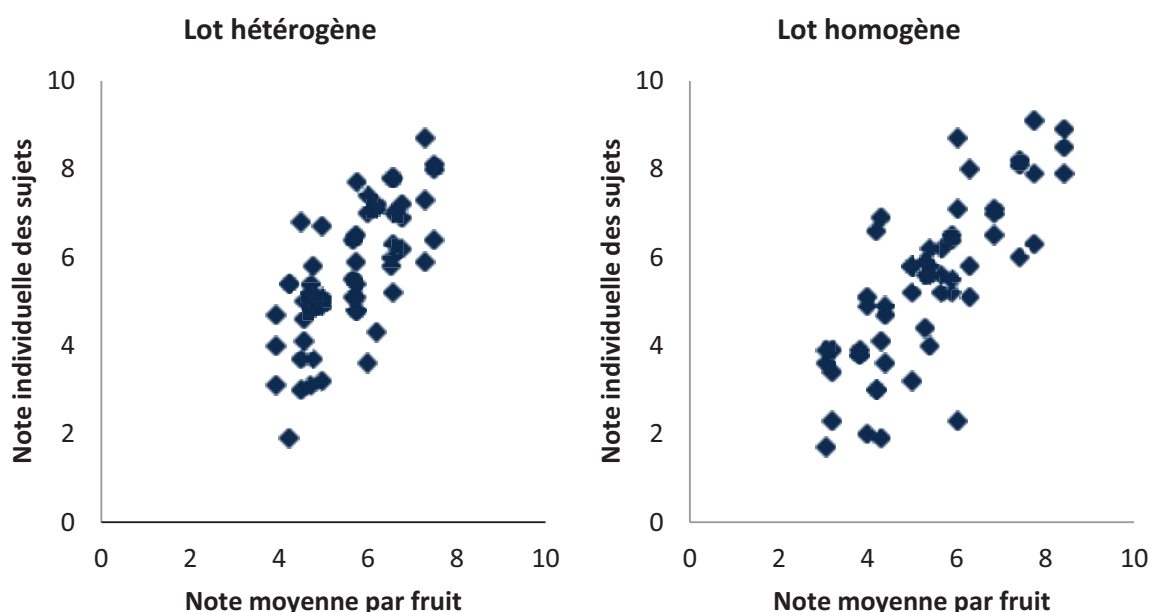


Figure 39. Note individuelles des sujets en fonction des notes moyennes des fruits pour la variété Ariane et descripteur fermeté pour le lot hétérogène (à gauche) et le lot homogène (à droite)

Le Tableau 33 et le Tableau 34 contiennent les résultats d'analyse de la variance avec le modèle « fruit » pour le profil des lots hétérogènes et le profil des lots homogènes respectivement. De même que pour l'échantillonnage destructif, les variations de la discrimination ($F_{\text{variété}}$) fluctuent de façon aléatoire (augmentation pour deux descripteurs et diminution pour cinq descripteurs). Cependant, la

valeur de discrimination ($F_{\text{variété}}$) pour le descripteur *acidité* augmente de façon importante ($3,26 < 21,64$). Cette augmentation est accompagnée d'une diminution de l'effet fruit ($22,34 > 11,33$), qui l'explique. Par contre, pour le descripteur *crissant*, la discrimination ($F_{\text{variété}}$) diminue ($13,86 > 3,42$). Pour rappel, le descripteur *crissant* n'avait pas permis la discrimination intra-lot lors de la mesure de performance intra-variétale (Chapitre I.2). La forte augmentation de l'effet fruit et de l'interaction sujet*variété peuvent expliquer cette diminution de l'effet variété.

1.4 Conclusion

L'homogénéisation instrumentale *a priori* des lots dégustés par le panel tend à réduire la variabilité de la réponse sensorielle et à améliorer les résultats de discrimination pour les trois variétés étudiées. Néanmoins, les résultats diffèrent entre l'homogénéisation par pénétrométrie et par sonométrie. La variabilité de la réponse sensorielle est réduite pour les variétés Ariane et Braeburn dans le cas d'utilisation de la pénétrométrie et pour la variété Pink Lady® dans le cas de l'utilisation de la mesure acoustique. Les réductions observées ne sont cependant pas significatives, ce qui peut être expliqué par les faibles corrélations observées entre les descripteurs sensoriels et les mesures instrumentales.

Les tendances de réduction de la variabilité observées pourraient être vérifiées dans le cadre d'une expérimentation incluant plus de mesures, augmentant la robustesse des résultats.

Tableau 33. Échantillonnage hétérogène par sonométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs

Hétérogène	Variété		Sujet		Sujet*Variété		Fruit		Erreur
	F _{2;46}	p-value	Chi ²	p-value	σ ²	p-value	Chi ²	p-value	
Croquant	21,01	2,89E-07***	15,49	8,30E-05***	0,68	4,09E-01	20,05	7,53E-06***	1.44
Fermeté	33,61	1,11E-09***	13,76	2,08E-04***	1,19	2,74E-01	11,43	7,24E-04***	1.35
Crissant	13,86	3,00E-06***	20,57	5,74E-06***	2,45	1,18E-01	3,40	6,52E-02	1.57
Jutosité	2,84	6,56E-02	0,77	3,82E-01	18,92	1,37E-05***	29,62	5,25E-08***	1.15
Fondant	16,56	3,18E-06***	14,49	1,41E-04***	4,95	2,60E-02*	29,91	4,53E-08***	1.43
Acidité	3,26	4,72E-02*	58,73	1,81E-14***	0,14	7,11E-01	22,34	2,29E-06***	1.30
Sucré	22,63	2,58E-09***	60,68	6,66E-15***	1,76	1,85E-01	2,22	1,37E-01	1.08

Niveau de significativité : 0,05* ; 0,01**, 0,001***

Tableau 34. Échantillonnage homogène par sonométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs

Homogène	Variété		Sujet		Sujet*Variété		Fruit		Erreur
	F _{2;46}	p-value	Chi ²	p-value	σ ²	p-value	Chi ²	p-value	
Croquant	15,03	8,38E-06***	22,91	1,70E-06***	0,03	8,60E-01	56,26	6,34E-14***	0.96
Fermeté	20,61	2,10E-07***	6,98	8,26E-03**	5,92	1,50E-02*	44,09	3,14E-11***	0.90
Crissant	3,42	3,98E-02*	10,25	1,36E-03**	7,76	5,34E-03**	35,14	3,07E-09***	0.84
Jutosité	4,14	2,11E-02*	1,86	1,72E-01	7,11	7,68E-03**	16,26	5,53E-05***	1.62
Fondant	14,70	8,83E-06***	8,60	3,36E-03**	4,21	4,01E-02*	31,11	2,43E-08***	1.66
Acidité	21,64	1,95E-07***	35,09	3,15E-09***	0,20	6,51E-01	11,33	7,63E-04***	1.80
Sucré	19,12	4,01E-08***	21,98	2,75E-06***	1,39	2,39E-01	0,95	3,30E-01	1.74

Niveau de significativité : 0,05* ; 0,01**, 0,001***

2. Le partage des fruits par plusieurs sujets permet de réduire la variabilité sensorielle

2.1 *Partage des fruits*

Nous avons choisi d'étudier l'impact du partage des fruits pour plusieurs sujets. En effet, cette méthode a été avancée par plusieurs auteurs comme limitant l'hétérogénéité des échantillons (Abbott *et al.*, 2004, Bhagwat *et al.*, 2004, Lateur *et al.*, 2001). La pomme est un fruit de taille moyenne qui peut aisément être coupé en huit morceaux, comme en témoignent le format des coupe-pommes. Cette division a d'ailleurs été utilisée par Lateur *et al.* (2001) dans le cadre du développement des aspects méthodologiques liés à l'utilisation de l'analyse sensorielle pour l'évaluation de la qualité gustative de pommes. Pour tester l'impact du nombre de sujets dégustant un même fruit, chaque fruit a été coupé en huit morceaux puis distribué à un ou plusieurs sujets. Les trois méthodes suivantes ont été appliquées :

- Un morceau à un sujet (1 sujet par fruit) ;
- Ou un morceau à quatre sujets (4 sujets par fruit) ;
- Ou un morceau à huit sujets (8 sujets par fruit).

Dans chaque situation les sujets ne recevaient qu'un seul morceau de pomme. Ainsi, les conditions de dégustation restaient les mêmes.

2.2 *Comparaison des méthodes de partage*

Afin d'évaluer l'efficacité de l'homogénéisation, les écart-types des résultats pour les trois modes de distribution ont été comparés ainsi que les résultats des ANOVAs standard (I-2-1). L'utilisation du modèle d'ANOVA standard se justifie par l'unicité des évaluations dans le cas du mode de présentation '1 sujet par fruit'.

2.3 *Résultats et discussion*

Les comparaisons des modes de distributions sont présentées deux-à-deux dans le Tableau 35, le Tableau 36 et le Tableau 37. Lorsque le nombre de sujets par fruit augmente, une diminution des écart-types serait attendue. Une diminution des écart-types est d'ailleurs observée entre le mode '1 sujet par fruit' et le mode '4 sujets par fruit', où les réductions des écarts-types varient entre 1% pour

le descripteur *croquant* et la variété Ariane et 21% pour le descripteur *fondant* et la variété Pink Lady® (Tableau 35). Par ailleurs, le partage des fruits pour plusieurs sujets peut également conduire à des résultats inverses comme par exemple pour le descripteur *fermeté* et la variété Pink Lady® pour lequel l'écart-type augmente de 14%. L'observation des différences entre les écart-types pour les modes '1 sujet par fruit' et '8 sujets par fruit' conduit à des résultats plus mitigés quant à la réduction par le partage des fruits (Tableau 36).

Les tests de comparaisons de variance indiquent que cette évolution est généralement non significative et dépendante des variétés. Les diminutions de variance sont significatives à 5% pour la variété Pink Lady® et le descripteur *fondant* dans les deux cas de partage (p-value = 3,61E-02* et p-value = 3,01E-02*) et le descripteur *croquant* dans le cas '8 sujets par fruit' (p-value = 9,17E-03***). Seules les variances du descripteur *fermeté*, pour la variété Pink Lady® et le mode '4 sujets par fruit', et pour la variété Braeburn et le '8 sujets par fruit', montrent une forte augmentation suite au partage.

Tableau 35. Écart-types de la réponse sensorielle pour les trois variétés de pommes, les cinq descripteurs et mode de présentation pour 1 et pour 4 sujets, et évolution des écart-types

Variété	Ariane			Braeburn			Pink Lady®			
	Nb sujets par fruit	1	4	Évol.	1	4	Évol.	1	4	Évol.
Croquant		1.72	1.70	-1%	1.67	1.55	-7%	1.06	1.00	-5%
Fermeté		1.79	1.63	-9%	1.61	1.63	1%	1.30	1.49	14%
Jutosité		1.55	1.56	0%	1.01	0.92	-9%	1.36	1.27	-7%
Fondant		2.14	1.83	-15%	2.01	1.69	-16%	1.40	1.10	-21%*
Acidité		1.88	1.83	-3%	2.16	2.08	-4%	2.03	1.67	-18%

Niveau de significativité : 0,05*

Tableau 36. Écart-types de la réponse sensorielle pour les trois variétés de pommes, les cinq descripteurs et mode de présentation pour 1 et pour 8 sujets, et évolution des écart-types

Variété	Ariane			Braeburn			Pink Lady®			
	Nb sujets par fruit	1	8	Évol.	1	8	Évol.	1	8	Évol.
Croquant		1.72	1.79	4%	1.67	1.43	-14%	1.06	0.78	-27%***
Fermeté		1.79	1.66	-7%	1.61	1.98	23%	1.30	1.35	3%
Jutosité		1.55	1.47	-5%	1.01	1.02	1%	1.36	1.42	4%
Fondant		2.14	2.00	-7%	2.01	1.77	-12%	1.40	1.09	-22%*
Acidité		1.88	1.92	2%	2.16	2.18	1%	2.03	1,94	-5%

Niveau de significativité : 0,05* ; 0,001***

Concernant la comparaison des modes '4 sujets par fruit' et '8 sujets par fruit' (Tableau 37), le mode '4 sujets par fruit' permet généralement d'obtenir des résultats moins variables. Cependant, les différences entre les écart-types ne sont pas significatives sauf pour la variété Pink Lady® et le descripteur *croquant* (p-value = 2,55E-02*). Les écart-types plus faibles généralement observés dans le cas du mode '4 sujets par fruit' peuvent être expliqués par la variabilité interne des fruits. En effet, la variabilité des morceaux opposés d'une pomme a été démontrée à plusieurs reprises (Dever *et al.*, 1995, Symoneaux *et al.*, 2002). Dans le cas du mode '4 sujets par fruit', les morceaux distribués étaient contigus. La probabilité qu'ils appartiennent à une même face était donc augmentée et la variabilité des morceaux proposés réduite. Dans le cas du mode '8 sujets par fruit', l'évaluation concernait forcément des morceaux des faces opposées.

Tableau 37. Écart-types de la réponse sensorielle pour les 5 descripteurs, 3 variétés et mode de présentation pour 4 et pour 8 sujets, et évolution des écart-types

Variété	Ariane			Braeburn			Pink Lady®			
	Nb sujets par fruit	4	8	Évol.	4	8	Évol.	4	8	Évol.
Croquant		1.70	1.79	6%	1.55	1.43	-8%	1.00	0.78	-23%*
Fermeté		1.63	1.66	1%	1.63	1.98	21%	1.49	1.35	-9%
Jutosité		1.56	1.47	-6%	0.92	1.02	11%	1.27	1.42	12%
Fondant		1.83	2.00	10%	1.69	1.77	5%	1.10	1.09	-1%
Acidité		1.83	1.92	5%	2.08	2.18	5%	1.67	1.94	16%

Niveau de significativité : 0,05*

Par ailleurs le partage des fruits entraîne généralement une augmentation de la statistique F liée à la discrimination (effet variété) (Tableau 38, Tableau 39 et Tableau 40). Cette augmentation de la discrimination s'explique par une diminution du terme d'interaction intervenant dans le calcul de cette statistique. Ce phénomène est facilement expliqué par le plus faible risque de désaccord entre les sujets lorsqu'ils dégustent le même fruit. Entre les modes '4 sujets par fruit' et '8 sujets par fruit', la discrimination des descripteurs *juteux* et *fondant* diminue. La variabilité interne des fruits peut là aussi expliquer cette observation. La détermination du nombre de sujets testant un même individu du lot semble un point critique et doit être réfléchi en fonction de la variabilité pouvant exister au sein de chaque fruit individuel.

Tableau 38. Mode '1 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs

Descripteurs	Facteur variété		Facteur sujet		Interaction sujet*variété	
	F _{2,45}	p-value	Chi ²	p-value	Chi ²	p-value
Croquant	27,92	1,94E-08***	3,24	7,17E-02	40,33	2,15E-10***
Fermeté	69,53	1,39E-11***	8,94	2,79E-03**	7,11	7,67E-03***
Jutosité	82,62	1,81E-12***	6,06	1,38E-02*	4,28	3,86E-02*
Fondant	60,99	6,28E-11***	4,89	2,70E-02*	5,77	1,63E-02*
Acidité	12,32	1,45E-04***	13,61	2,25E-04***	4,64	3,12E-02*

Niveau de significativité : 0,05* ; 0,01** , 0,001***

Tableau 39. Mode '4 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs

Descripteurs	Facteur variété		Facteur sujet		Interaction sujet*variété	
	F _{2,45}	p-value	Chi ²	p-value	Chi ²	p-value
Croquant	48,84	7,45E-10***	9,88	1,67E-03**	6,14	1,32E-02*
Fermeté	155,17	1,85E-38***	60,51	7,33E-15***	0,15	6,97E-01
Jutosité	56,59	1,21E-12***	1,71	1,91E-01	28,36	1,00E-07***
Fondant	60,85	3,93E-13***	1,73	1,89E-01	31,35	2,15E-08***
Acidité	36,78	6,56E-14***	59,25	1,39E-14***	0,96	3,27E-01

Niveau de significativité : 0,05* ; 0,01** , 0,001***

Tableau 40. Mode '8 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs

Descripteurs	Facteur variété		Facteur sujet		Interaction sujet*variété	
	F _{2,45}	p-value	Chi ²	p-value	Chi ²	p-value
Croquant	31,29	4,78E-09***	0,79	3,75E-01	31,99	1,55E-08***
Fermeté	126,41	7,13E-34***	44,94	2,03E-11***	0,00	1,00E+00
Jutosité	92,09	4,88E-13***	8,62	3,33E-03***	5,64	1,75E-02*
Fondant	89,90	6,54E-13***	5,38	2,04E-02*	6,82	9,03E-03***
Acidité	39,83	8,24E-15***	65,63	5,55E-16***	0,59	4,44E-01

Niveau de significativité : 0,05* ; 0,01** , 0,001***

L'étude des composantes de la variance pour les méthodes '4 sujets par fruit' et '8 sujets par fruit' indique qu'il y a peu de différences de répartitions des composantes entre les deux méthodes

(Tableau 41). Ces résultats peuvent être comparés à ceux de Guinness *et al.* (2009). Dans cette étude, les résultats des évaluations sensorielles de fraises et de purées de fraise ont été comparés en termes de distribution des composantes de la variance. Les sujets ont été identifiés comme participant de façon conséquente à la variabilité totale, notamment pour trois descripteurs (*odeur herbe*, *sucré* et *amer*). En parallèle, la forte contribution des fruits individuels a été soulignée (e.g. plus de 60% pour le descripteur *sucré*), comparée à la contribution des échantillons de purée (30 à 40% quel que soit le descripteur). Similairement à cette étude, dans notre cas, les sujets contribuent fortement aux variations de la réponse. La contribution des sujets dépend des descripteurs, allant de 9% à 69% (les deux méthodes confondues). Dans les deux cas, la contribution des sujets est plus forte pour les descripteurs *fermeté* et *acidité*. Cette part de variation due aux sujets reflète les différences d'utilisation de l'échelle en termes de moyenne de notation (effet niveau). Comme mentionné précédemment, cet effet de niveau est courant en évaluation sensorielle. En parallèle, la contribution du fruit est proche et dépasse même en général les 20% (excepté pour le descripteur *jutosité* et la méthode '4 sujets par fruit').

Contrairement au constat fait par Guinness *et al.* (2009), qui ont observé des contributions à la variance généralement plus faible pour les échantillons de purée que pour les fruits individuels, pour les cinq descripteurs étudiés, seule une diminution de la contribution des fruits pour le descripteur *fermeté* est observée. Pour les quatre autres descripteurs, les contributions sont similaires (*croquant*, *fondant* et *acidité*) ou supérieure (*jutosité*). Cependant, dans l'étude de Guinness *et al.* (2009), la diminution des contributions peut s'expliquer par un processus d'homogénéisation plus drastique que le simple partage des fruits puisqu'il consistait à passer de fruits à une purée. Cette solution n'est pas idéale car la purée ne permet pas l'évaluation de la texture des fruits et implique des modifications des odeurs et flaveur des produits étudiés par rapport à la matrice initiale.

2.4 Conclusion

La dégustation des fruits par plusieurs sujets permet de réduire la variabilité des résultats sensoriels dans la plupart des cas. De plus, le partage des fruits permet une augmentation de la discrimination des variétés. Néanmoins, le partage des fruits permet aussi la prise en compte de la variabilité intra-lot qui entraîne une chute de la discrimination (Chapitre II.1). Par ailleurs, le partage des fruits a montré un intérêt indirect au cours de l'entraînement. Il a été remarqué que ce mode de distribution permettait de mettre les sujets plus en confiance par rapport à leur performance. En effet, par exemple lors d'une évaluation en consensus, si chaque sujet reçoit une pomme différente, leurs notes pour cette variété peuvent différer de façon importante. Lorsque plusieurs sujets évaluent la

même pomme, ils sont généralement en accord. Lorsque les notes de l'ensemble des sujets du panel diffèrent et que chaque sous-panel ayant dégusté la même pomme est en accord, les différences entre les notes peuvent alors être attribuées aux différences entre les pommes. Enfin, l'application du modèle proposé au Chapitre II.1 n'est possible que si chaque fruit est dégusté par au moins deux sujets, renforçant à nouveau la pertinence d'un tel mode de présentation.

Tableau 41. Composantes de la variance et , entre parenthèses, leur pourcentage de la variation totale avec méthode '4 sujets par fruit' et la méthode '8 sujets par fruit' pour les cinq descripteurs

Descripteurs	'4 sujets par fruit'					'8 sujets par fruit'				
	$\sigma_{\text{variété}}$	σ_{sujet}	$\sigma_{\text{sujet} \times \text{variété}}$	σ_{fruit}	σ_{erreur}	$\sigma_{\text{variété}}$	σ_{sujet}	$\sigma_{\text{sujet} \times \text{variété}}$	σ_{fruit}	σ_{erreur}
Croquant	48,39	13,44 (48)	8,94 (32)	5,83 (21)	23,4	52,21	2,18 (09)	16,25 (67)	5,75 (24)	23,61
Fermeté	59,42	16,22 (64)	2,23 (09)	6,86 (27)	15,26	56,59	14,47 (62)	1,29 (06)	7,60 (03)	20,04
Jutosité	66,42	3,62 (26)	8,97 (65)	1,30 (09)	19,69	63,77	9,00 (46)	6,66 (34)	3,95 (20)	16,61
Fondant	67,84	3,98 (21)	9,30 (50)	5,42 (29)	13,45	65,33	6,42 (32)	7,78 (38)	6,02 (30)	14,44
Acidité	25,20	29,58 (69)	5,84 (14)	7,69 (18)	31,69	25,47	31,18 (69)	5,66 (13)	8,28 (18)	29,41

3. Conclusion générale

Les essais de réduction de la variabilité sont encourageants mais n'ont pas donné de résultats significatifs. L'homogénéisation par pénétrométrie et le partage des fruits semblent pertinents au vu des tendances de diminution de la variabilité de la réponse sensorielle. Concernant l'homogénéisation par mesures instrumentales, le choix de son utilisation devrait être fait en tenant compte du gain effectif sur la variabilité et du coût de cette procédure. Le partage des fruits pour plusieurs sujets permet de réduire la variabilité de la réponse sensorielle (tendance). Cette stratégie impose peu de contraintes dans le cas de la pomme et permet, qui plus est, l'utilisation du modèle « fruit » décrit au Chapitre II.1.

Partie V. Conclusion et perspectives

Ce travail de thèse est né du challenge que représente la variabilité biologique des fruits lors de l'évaluation de la qualité et plus particulièrement de celle des caractéristiques sensorielles. Deux approches ont été explorées pour répondre à ce constat : la mesure de la variabilité de la réponse sensorielle avec le choix des modèles d'analyse de la variance à utiliser et la réduction de la variabilité de la réponse sensorielle par des méthodes d'homogénéisation instrumentale ou par le partage des fruits.

Avant d'aborder ces deux approches, une étude des performances du panel recruté et entraîné dans le cadre de cette thèse a été menée. Dans un premier temps, l'étude de la capacité de discrimination inter-variétale (inter-lot) du panel a été réalisée, suivie, dans un second temps, de l'étude de la capacité de discrimination intra-variétale (intra-lot) du panel.

La capacité du panel à percevoir des différences pour les sept descripteurs étudiés a été démontrée. Pour les descripteurs *croquant*, *crissant* et *jutosité*, la variété Pink Lady® reçoit les notes les plus élevées, suivie de la variété Braeburn puis de la variété Ariane. Par ailleurs, les descripteurs *fermeté* et *fondant* permettent de différencier la variété Pink lady® des deux autres variétés moins fermes et plus fondantes. Concernant les saveurs, la variété Ariane est la plus acide et la plus sucrée. La variété Braeburn est, quant à elle, la moins acide et la moins sucrée. La variété Pink Lady® se place en position intermédiaire pour ces deux descripteurs de saveur.

Le panel a également démontré sa capacité à discriminer des fruits d'une même variété (un même lot) ayant des niveaux de fermeté acoustique différents pour tous les descripteurs étudiés, excepté

crissant. En général, les pommes de fermeté acoustique basse étaient perçues moins croquantes, fermes, juteuses, acides et sucrées et plus fondantes que les pommes de fermeté acoustique haute. Néanmoins, des différences de discrimination entre les fruits des différents niveaux de fermeté acoustique sont observées en fonction des variétés. Notamment, la discrimination au sein de la variété Braeburn était plus importante qu'au sein des deux autres variétés. Ces différences peuvent s'expliquer par une dépendance de la mesure acoustique vis-à-vis des variétés, ou par les interactions sensorielles influençant la réponse des sujets.

Une fois les capacités de discrimination du panel démontrées tant au niveau inter-variétal qu'intra-variétal, les deux approches mentionnées précédemment ont été abordées.

La première approche avait pour objectif de tester des modèles d'analyse de la variance pour le traitement de la réponse sensorielle. Ce travail a été réalisé en deux temps.

D'abord, un modèle permettant la prise en compte de la variabilité intra-lot a été étudié par l'ajout d'un facteur fruit aléatoire autorisé par la dégustation de chaque fruit par plusieurs sujets, donnant naissance au modèle « fruit ». L'intégration de ce nouveau facteur a permis de démontrer l'importance de la variabilité intra-lot qui se révèle souvent supérieure à la variabilité attribuée aux sujets (facteur sujet et interaction sujet*variété). L'introduction du facteur fruit induit des modifications quant au test de l'effet variété. Avec le modèle standard, cet effet est testé contre l'interaction sujet*variété, alors que dans le cas présent il est testé contre une combinaison de l'interaction sujet*variété et du facteur fruit. Cette différence implique notamment une diminution de l'effet variété avec l'utilisation du modèle « fruit », comparativement à l'utilisation du modèle standard. Elle s'explique par l'augmentation du bruit dans le ratio signal/bruit permettant de tester la discrimination. Par ailleurs, par le biais d'une simulation de données, l'impact de la prise en compte de la variabilité intra-lot lorsque celle-ci est particulièrement forte a été démontré. La prise en compte de la variabilité intra-lot apporte des résultats plus conservateurs quant à la conclusion sur les différences entre les produits étudiés.

Puis, une étude des différences entre les sujets du panel a mis en évidence qu'une grande partie de l'interaction sujet*variété pouvait être attribuée aux différences de dispersion entre les sujets plus qu'au désaccord pur entre ceux-ci. Afin de remédier à l'interprétation erronée de ce terme d'interaction, Brockhoff et Skovgaard (1994) ont proposé un modèle baptisé « *assessor model* ». Plus récemment, un modèle étendu a été proposé et comprend un terme d'interaction, reflet du désaccord pur, et utilise les notes moyennes centrées des produits comme co-variables pour intégrer l'effet de dispersion (Brockhoff *et al.*, 2012). Une approche similaire a été appliquée au modèle

« fruit » afin de ne prendre en compte que le désaccord pur dans le terme d'interaction. Les données ont été corrigées pour la dispersion, puis le modèle « fruit » a été utilisé. Les valeurs des effets variétés sont ici augmentées, comparativement à l'utilisation du modèle « fruit » seul. Ce phénomène est expliqué par la diminution de la valeur de la partie bruit du ratio signal/bruit permettant de tester l'effet variété. L'approche « assessor » semble donc pertinente afin de tester l'effet produit contre ce qui représente vraiment le bruit de cette mesure. La vérification de l'impact de la transformation des données pour l'approche « assessor » a démontré une modification du classement des produits dans les cas présentant initialement un désaccord. Il sera nécessaire de s'assurer que le modèle proposé intégrant effet fruit et distinguant le désaccord pur et l'effet de dispersion ne soustrait que la variabilité due aux différences de dispersion de l'interaction.

L'étude des modèles d'analyse de la variance a, par ailleurs, soulevé des interrogations quant à la détermination des tests les plus appropriés en fonction de la significativité des effets aléatoires testés. Notamment, dans le modèle standard, lorsque l'interaction sujet*produit n'est pas significative, son carré moyen attendu reflète alors une mesure de l'erreur résiduelle. Dans ce cas, choisir la variabilité la plus élevée pour tester l'effet permettra l'obtention de résultats plus conservateurs et l'effet produit devrait être testé contre le terme d'erreur. Cette remarque peut s'étendre au modèle « fruit » : dans le cas d'un effet fruit non significatif et d'une interaction significative, l'effet produit devrait alors être testé contre le terme d'interaction seul ; et dans le cas de la non-significativité de l'effet fruit et du terme d'interaction, l'effet produit serait alors testé contre le terme d'erreur.

Le modèle d'analyse de la variance proposé répond à des questions présentes à l'esprit des utilisateurs de l'évaluation sensorielle. Le modèle « fruit » devrait selon nous se substituer au modèle standard pour toute étude de produits susceptibles de présenter une variabilité intra-lot. Cet outil pourra éviter toutes conclusions erronées quant à la perception de différences entre les produits. D'un point de vue pratique et étant donné les contraintes expérimentales, plusieurs points restent encore à étudier et notamment au regard de la robustesse du modèle : le nombre de répétitions nécessaire, le nombre de sujets du panel et le nombre de sujets dégustant un même individu. La question de la construction du plan d'expérience vient s'ajouter à ces préoccupations. Les sujets dégustant un même individu du lot doivent être randomisés afin d'éviter la confusion de la variabilité intra-lot et d'un éventuel effet de niveau lié à la constitution des groupes de sujets.

La deuxième approche concernait l'évaluation de l'efficacité d'une réduction de la variabilité *a priori* des lots sur la réduction de la variabilité de la réponse sensorielle. Deux méthodes ont été évaluées : l'échantillonnage instrumental et le partage des fruits.

Concernant l'échantillonnage instrumental, deux mesures ont été testées : une mesure destructive, la pénétrométrie et une mesure non destructive : la sonométrie. La réduction de la variabilité de la réponse par l'homogénéisation instrumentale *a priori* a été démontrée pour certaines variétés et certains descripteurs. Les réductions observées ne sont cependant pas significatives. Des tendances de réduction de la variabilité sensorielle sont observées pour l'échantillonnage par pénétrométrie des fruits issus des variétés Ariane et Braeburn, ainsi que pour l'échantillonnage par sonométrie de la variété Pink Lady®. L'impact de l'homogénéisation *a priori* sur les résultats d'analyse de la variance ne correspond pas à celui attendu, qui était une amélioration de la discrimination. Les faibles corrélations observées entre les descripteurs sensoriels et les mesures instrumentales pourraient expliquer l'impact limité de la réduction *a priori* de la variabilité des lots dégustés sur la variabilité de la réponse sensorielle, de même que la complexité de certains descripteurs.

L'homogénéisation par pénétrométrie a montré des tendances à la réduction de la variabilité de la réponse sensorielle. Cette méthode pourrait alors être utilisée dans le cadre d'une évaluation de performance afin de minimiser l'intervention de la variabilité intra-lot. Cependant, l'échantillonnage est une procédure longue et nous ne recommandons pas son utilisation en routine pour la réalisation de profil. L'homogénéisation par sonométrie n'a pas procuré les résultats attendus. Cette méthode ne semble donc pas pertinente. Son avantage consiste dans sa rapidité d'exécution et sa nature non destructive. D'autres méthodes (e.g. spectroscopie laser résolue dans l'espace) présentant les mêmes avantages pourraient être étudiées dans le cadre de la réduction *a priori* de la variabilité.

Concernant le partage des fruits, des tendances de diminution de la variabilité de la réponse sensorielle sont observées lorsque les fruits sont partagés. Les résultats dépendent de la stratégie de partage adoptée. Dans le cas de la pomme, le partage d'un fruit pour quatre sujets semble plus approprié. La distribution d'un fruit entier, dans le cas du partage pour huit sujets semble en effet introduire de la variabilité pouvant être expliquée par les différences perceptibles entre la face ensoleillée et la face non ensoleillée.

Le partage des fruits a démontré des tendances de réduction de la variabilité de la réponse sensorielle. Il peut avoir des conséquences lors des phases d'entraînement des sujets. En effet, lors du travail en consensus, la dégustation d'un fruit par plusieurs sujets donnera de meilleurs résultats que si les sujets dégustaient tous des fruits différents. Par ailleurs, ce mode de distribution est essentiel pour appliquer la méthode d'analyse tenant compte de la variabilité fruit.

Finalement, les méthodologies proposées suite à ce travail sont applicables pour l'évaluation de la qualité sensorielle de tout produit variable et répondent à un problème concret rencontré par de nombreux utilisateurs de l'évaluation sensorielle. Elles devront être adaptées aux objectifs de

l'expérimentateur. L'application du modèle « fruit » pour une meilleure analyse de la qualité sensorielle nécessite le partage des individus du lot pour plusieurs sujets. Le nombre de sujets dégustant un même individu pourra être adapté au besoin de représentativité de l'étude et au nombre de répétitions pouvant être effectuées. De même, le choix de l'homogénéisation sera conditionné par le besoin ou non de représentativité. L'homogénéisation pourra être utilisée dans le cadre de mesure de performance ou dans une étude visant à identifier de façon sûre l'influence de certains paramètres. Dans ce dernier cas, suite à l'identification des facteurs influents, la mesure pourrait être répétée sur un échantillon représentatif, non homogénéisé, afin de valider les résultats. Enfin, une attention particulière devra être portée au transfert des méthodes statistiques vers les utilisateurs de l'évaluation sensorielle.

Les données de ce travail de thèse sont accessibles. Contacter Ronan Symoneaux pour plus d'information : r.symoneaux@groupe-esa.com

Références

- A**aby, K., Haffner, K., & Skrede, G. (2002). Aroma quality of Gravenstein apples influenced by regular and controlled atmosphere storage. *Lebensmittel-Wissenschaft und-Technologie*, 35 (3), 254-259.
- Abbott, J.A., Affeldt, H.A., & Liljedahl, L.A. (1992). Firmness measurement of stored 'Delicious' apples by sensory methods, Magness-Taylor, and sonic transmission. *Journal of the American Society for Horticultural Science*, 117 (4), 590-595.
- Abbott, J.A., Saftner, R.A., Gross, K.C., Vinyard, B.T., & Janick, J. (2004). Consumer evaluation and quality measurement of fresh-cut slices of 'Fuji', 'Golden Delicious', 'GoldRush', and 'Granny Smith' apples. *Postharvest biology and technology*, 33 (2), 127-140.
- Ackermann, J., Fischer, M., & Amado, R. (1992). Changes in sugars, acids, and amino acids during ripening and storage of apples (cv. Glockenapfel). *Journal of agricultural and food chemistry*, 40 (7), 1131-1134.
- AFNOR NF X 06-004 (1991). *Vocabulaire du contrôle de la qualité*. International Organization for Standardization,
- Aguayo, E., Requejo-Jackman, C., Stanley, R., & Woolf, A. (2010). Effects of calcium ascorbate treatments and storage atmosphere on antioxidant activity and quality of fresh-cut apple slices. *Postharvest biology and technology*, 57 (1), 52-60.
- Allan-Wojtas, P., Sanford, K., McRae, K., & Carbyn, S. (2003). An integrated microstructural and sensory approach to describe apple texture. *Journal of the American Society for Horticultural Science*, 128 (3), 381-390.
- B**arreiro, P., Ortiz, C., Altisent, M.R., Schotte, S., Andani, Z., Wakeling, I., & Beyt, P.K. (1998). Comparison between sensory and instrumental measurements for mealiness assessment in apples. A collaborative test. *Journal of Texture Studies*, 29 (5), 509-525.
- Bartoshuk, L.M. (2000). Comparing sensory experiences across individuals: recent psychophysical advances illuminate genetic variation in taste perception. *Chemical senses*, 25 (4), 447.
- Bartoshuk, L.M., Duffy, V.B., & Miller, I.J. (1994). PTC/PROP tasting: Anatomy, psychophysics, and sex effects. *Physiology & Behavior*, 56 (6), 1165-1171.
- Barylko-Pikielna, N., Zawirska-Wojtasiak, R., Kornelson, C., & Rothe, M. (1992). Interlaboratory reproducibility of sensory profiling results of margarines. In, *3rd Wartburg Aroma Symposium*. Eisenach.
- Bhagwat, A.A., Saftner, R.A., & Abbott, J.A. (2004). Evaluation of wash treatments for survival of foodborne pathogens and maintenance of quality characteristics of fresh-cut apple slices. *Food microbiology*, 21 (3), 319-326.

- Bitnes, J., Martens, H., Ueland, Ø., & Martens, M. (2007a). Longitudinal study of taste identification of sensory panellists: Effect of Ageing, Experience and Exposure. *Food Quality and Preference*, 18 (2), 230-241.
- Bitnes, J., RoDbotten, M., Lea, P., Ueland, Y., & Martens, M. (2007b). Effect of product knowledge on profiling performance comparing various sensory laboratories. *Journal of Sensory Studies*, 22 (1), 66-80.
- Blakeslee, A.F., & Salmon, T.N. (1935). Genetics of sensory thresholds: individual taste reactions for different substances. *Proceedings of the National Academy of Sciences of the United States of America*, 21 (2), 84.
- Bobelyn, E., Serban, A.S., Nicu, M., Lammertyn, J., Nicolai, B.M., & Saeys, W. (2010). Postharvest quality of apple predicted by NIR-spectroscopy: Study of the effect of biological variability on spectra and model performance. *Postharvest biology and technology*, 55 (3), 133-143.
- Bourne, M.C., (2002). *Food texture and viscosity: concept and measurement*. Academic Press.
- Brandt, M.A., Skinner, E.Z., & Coleman, J.A. (1963). Texture profile method. *Journal of Food Science*, 28 (4), 404-409.
- Brockhoff, P.B. (2003). Statistical testing of individual differences in sensory profiling. *Food Quality and Preference*, 14 (5-6), 425-434.
- Brockhoff, P.B., Schlich, P., & Skovgaard, I.M. (2012). Accounting for scaling differences in sensory profile data: improved mixed model analysis of variance. In, *manuscript intended for Food Quality and Preference*.
- Brockhoff, P.B., & Skovgaard, I.M. (1994). Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference*, 5 (3), 215-224.
- Brookfield, P.L., Nicoll, S., Gunson, F.A., Harker, F.R., & Wohlers, M. (2011). Sensory evaluation by small postharvest teams and the relationship with instrumental measurements of apple texture. *Postharvest biology and technology*, 59 (2), 179-186.
- Broom, F.D., Smith, G.S., Miles, D.B., & Green, T.G.A. (1998). Within and between tree variability in fruit characteristics associated with bitter pit incidence of 'Braeburn' apple. *Journal of horticultural science & biotechnology*, 73 (4), 555-561.
- Brossard, C., Martin, N., Salles, C., & Guichard, E., (2012). *Influences respectives de la texture, de la saveur et de l'arôme vues par un dégustateur*, Texture et flaveur des aliments, Educagri Editions, pp. 263-292.
- Cairncross, S.E., & Sjostrom, L.B. (1950). Flavor profiles—A new approach to flavor problems. *Food Technology*, 4 (8), 308.
- Carbonell, L., Izquierdo, L., & Carbonell, I. (2007). Sensory analysis of Spanish mandarin juices. Selection of attributes and panel performance. *Food Quality and Preference*, 18 (2), 329-341.
- Chen, H., & DeBaerdemaeker, J. (1993). Effect of Apple Shape on Acoustic Measurements of Firmness. *Journal of Agricultural Engineering Research*, 56 (3), 253-266.

Couronne, T. (1997). A study of assessors' performance using graphical methods. *Food Quality and Preference*, 8 (5-6), 359-365.

Dahl, T., & Næs, T. (2004). Outlier and group detection in sensory panels using hierarchical cluster analysis with the Procrustes distance. *Food Quality and Preference*, 15 (3), 195-208.

Dahl, T., & Næs, T. (2009). Identifying outlying assessors in sensory profiling using fuzzy clustering and multi-block methodology. *Food Quality and Preference*, 20 (4), 287-294.

Daillant-Spinnler, B., MacFie, H.J.H., Beyts, P.K., & Hedderley, D. (1996). Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference*, 7 (2), 113-126.

Dairou, V., & Sieffermann, J.M. (2002). A comparison of 14 jams characterized by conventional profile and a quick original method, the flash profile. *Journal of Food Science*, 67 (2), 826-834.

De Belie, N., Schotte, S., Coucke, P., & De Baerdemaeker, J. (2000). Development of an automated monitoring device to quantify changes in firmness of apples during storage. *Postharvest biology and technology*, 18 (1), 1-8.

De Ketelaere, B., Stulens, J., Lammertyn, J., Cuong, N.V., & De Baerdemaeker, J. (2006). A methodological approach for the identification and quantification of sources of biological variance in postharvest research. *Postharvest biology and technology*, 39 (1), 1-9.

DeEll, J.R., & Prange, R.K. (1992). Postharvest quality and sensory attributes of organically and conventionally grown apples. *HortScience*, 27 (10), 1096-1099.

Depled, F., (2009). *Évaluation sensorielle: manuel méthodologique*. Lavoisier.Paris.

Dever, M.C., Cliff, M.A., & Hall, J.W. (1995). Analysis of variation and multivariate relationships among analytical and sensory characteristics in whole apple evaluation. *Journal of the Science of Food and Agriculture*, 69 (3), 329-338.

Dijksterhuis, G. (1995). Assessing panel consonance. *Food Quality and Preference*, 6 (1), 7-14.

Duizer, L. (2001). A review of acoustic research for studying the sensory perception of crisp, crunchy and crackly textures. *Trends in Food Science & Technology*, 12 (1), 17-24.

Duprat, F., Grotte, M., Loonis, D., & Pietri, E. (2000). Etude de la possibilité de mesurer simultanément la fermeté de la chair et de l'épiderme des pommes. *Sciences des aliments*, 20, 253-264.

Etaio, I., Albisu, M., Ojeda, M., Gil, P.F., Salmerón, J., & Pérez Elortondo, F.J. (2010). Sensory quality control for food certification: A case study on wine. Panel training and qualification, method validation and monitoring. *Food Control*, 21 (4), 542-548.

Evin, F., & Siekierski, E. (2002). Sensory evaluation of heating and air conditioning systems. *Energy and Buildings*, 34 (6), 647-651.

Feliciano, R.P., Antunes, C., Ramos, A., Serra, A.T., Figueira, M.E., Duarte, C.M.M., Carvalho, A., & Bronze, M.R. (2010). Characterization of traditional and exotic apple varieties from Portugal. Part 1-Nutritional, phytochemical and sensory evaluation. *Journal of Functional Foods*, (2), 35-45.

Findlay, C.J., Castura, J.C., & Lesschaeve, I. (2007). Feedback calibration: A training method for descriptive panels. *Food Quality and Preference*, 18 (2), 321-328.

Findlay, C.J., Castura, J.C., Schlich, P., & Lesschaeve, I. (2006). Use of feedback calibration to reduce the training time for wine panels. *Food Quality and Preference*, 17 (3-4), 266-276.

Finney, E. (1971). Dynamic elastic properties and sensory quality of apple fruit. *Journal of Texture Studies*, 2 (1), 62-74.

FranceAgrimer, (2012). *La pomme - Bilan de campagne 2010/2011*.

Frétillet, J.-P. (2010). Les petits secret des pommes. *ça m'intéresse*, 356, 88-90.

García-Ramos, F.J., Valero, C., Homer, I., Ortiz-Cañavate, J., & Ruiz-Altisent, M. (2005). Non-destructive fruit firmness sensors: a review. *Spanish Journal of Agricultural Research*, 3 (1), 61-73.

Gunness, P., Kravchuk, O., Nottingham, S.M., D'Arcy, B.R., & Gidley, M.J. (2009). Sensory analysis of individual strawberry fruit and comparison with instrumental analysis. *Postharvest biology and technology*, 52 (2), 164-172.

Hampson, C.R., Quamme, H.A., Hall, J.W., MacDonald, R.A., King, M.C., & Cliff, M.A. (2000). Sensory evaluation as a selection tool in apple breeding. *Euphytica*, 111 (2), 79-90.

Harker, F.R., Gunson, F.A., Brookfield, P.L., & White, A. (2002a). An apple a day: the influence of memory on consumer judgment of quality. *Food Quality and Preference*, 13 (3), 173-179.

Harker, F.R., Kupferman, E.M., Marin, A.B., Gunson, F.A., & Triggs, C.M. (2008). Eating quality standards for apples based on consumer preferences. *Postharvest biology and technology*, 50 (1), 70-78.

Harker, F.R., Maindonald, J., Murray, S.H., Gunson, F.A., Hallett, I.C., & Walker, S.B. (2002b). Sensory interpretation of instrumental measurements. 1: texture of apple fruit. *Postharvest biology and technology*, 24 (3), 225-239.

Harker, F.R., Maindonald, J.H., & Jackson, P.J. (1996). Penetrometer measurement of apple and kiwifruit firmness: operator and instrument differences. *Journal of the American Society for Horticultural Science*, 121 (5), 927-936.

- Harker, F.R., Marsh, K.B., Young, H., Murray, S.H., Gunson, F.A., & Walker, S.B. (2002c). Sensory interpretation of instrumental measurements. 2: Sweet and acid taste of apple fruit. *Postharvest biology and technology*, 24 (3), 241-250.
- Heintz, C.M., & Kader, A.A. (1983). Procedures for the sensory evaluation of horticultural crops [Quality, color and texture]. *HortScience*, 18 (1), 18-22.
- Hertog, M.L. (2002). The impact of biological variation on postharvest population dynamics. *Postharvest biology and technology*, 26 (3), 253-263.
- Hertog, M.L., Lammertyn, J., Desmet, M., & Scheerlinck, N. (2004). The impact of biological variation on postharvest behaviour of tomato fruit. *Postharvest biology and technology*, 34 (3), 271-284.
- Husson, F., Bocquet, V., & Pages, J. (2004). Use of confidence ellipses in a PCA applied to sensory analysis application to the comparison of monovarietal ciders. *Journal of Sensory Studies*, 19 (6), 510-518.
- Husson, F., Lê, S., & Pages, J. (2005). Confidence ellipse for the sensory profiles obtained by principal component analysis. *Food Quality and Preference*, 16 (3), 245-250.
- Huynh, M.-N., Martin, C., Cadoret, M., & Husson, F. (2012). Construction of confidence area in sensometry. In *11th Sensometrics*,
- Ioannides, Y., Howarth, M.S., Raithatha, C., Defernez, M., Kemsley, E.K., & Smith, A.C. (2007). Texture analysis of Red Delicious fruit: Towards multiple measurements on individual fruit. *Food Quality and Preference*, 18 (6), 825-833.
- ISO 874 (1980). *Fruit et légume en l'état - Echantillonnage. Echantillonnage et contrôle des produits alimentaires*. International Organization for Standardization, pp. 677-680.
- ISO 5492:2008 (1992). *Analyse sensorielle - Vocabulaire*. International Organization for Standardization, pp. 9-30.
- ISO 8586-1 (1993). *Analyse sensorielle - Lignes directrices générales pour la sélection, l'entraînement et le contrôle des sujets - Partie 1 : sujets qualifiés*. International Organization for Standardization, pp 59-86.
- ISO 11035 (1994). *Analyse sensorielle - Recherche et sélection de descripteurs pour l'élaboration d'un profil sensoriel, par approche multidimensionnelle*. International Organization for Standardization, pp 521-546.
- ISO 13299 (2003). *Analyse sensorielle - Méthodologie - Directives générales pour l'établissement d'un profil sensoriel*. International Organization for Standardization, pp 475-506.
- ISO 5495 (2005a). *Analyse sensorielle - Méthodologie - Essai de comparaison par paires*. International Organization for Standardization, pp 321-343.
- ISO 17025 (2005b). *Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais*. International Organization for Standardization,
- ISO 9000 (2005c). *Systèmes de management de la qualité - Principes essentiels et vocabulaire*. International Organization for Standardization,

ISO 8589:2007 (2007a). *Analyse sensorielle - Directives générales pour la conception de locaux destinés à l'analyse*. International Organization for Standardization, pp 133-145.

ISO / CEI GUIDE 99 (2007b). *Vocabulaire International de Métrologie – Concepts fondamentaux et généraux et termes associés (VIM)*

Jackman, R.L., & Stanley, D.W. (1995). Perspectives in the textural evaluation of plant foods. *Trends in Food Science & Technology*, 6 (6), 187-194.

Jaeger, S.R., Andani, Z., Wakeling, I.N., & MacFie, H.J.H. (1998). Consumer preferences for fresh and aged apples: a cross-cultural comparison. *Food Quality and Preference*, 9 (5), 355-366.

Jobling, J.J., & McGlasson, W.B. (1995). A comparison of ethylene production, maturity and controlled atmosphere storage life of Gala, Fuji and Lady Williams apples (*Malus domestica*, Borkh.). *Postharvest biology and technology*, 6 (3-4), 209-218.

Karlsen, A.M., Aaby, K., Sivertsen, H., Baardseth, P., & Ellekjær, M.R. (1999). Instrumental and sensory analysis of fresh Norwegian and imported apples. *Food Quality and Preference*, 10 (4-5), 305-314.

Khan, A.A., & Vincent, J.F.V. (1993). Anisotropy in the fracture properties of apple flesh as investigated by crack-opening tests. *Journal of Materials Science*, 28 (1), 45-51.

King, G.J., Maliepaard, C., Lynn, J.R., Alston, F.H., Durel, C.E., Evans, K.M., Griffon, B., Laurens, F., Manganaris, A.G., & Schrevens, E. (2000). Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). *TAG Theoretical and Applied Genetics*, 100 (7), 1074-1084.

King, M.C., Hall, J., & Cliff, M.A. (2001). A comparison of methods for evaluating the performance of a trained sensory panel. *Journal of Sensory Studies*, 16 (6), 567-582.

Konopacka, D., & Plochanski, W.J. (2004). Effect of storage conditions on the relationship between apple firmness and texture acceptability. *Postharvest biology and technology*, 32 (2), 205-211.

Kühn, B.F., & Thybo, A. (2001). Sensory quality of scab-resistant apple cultivars. *Postharvest biology and technology*, (23), 41-50.

Kuznetsova, A., Christensen, R.H.B., & Brockhoff, P.B., (2012). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 1.0-2.

Labbe, D., Rytz, A., & Hugi, A. (2004). Training is a critical step to obtain reliable product profiles in a real food industry context. *Food Quality and Preference*, 15 (4), 341-348.

Lateur, M., Planchon, V., & Moons, E. (2001). Évaluation par l'analyse sensorielle des qualités organoleptiques d'anciennes variétés de pommes. *Biotechnol. Agron. Soc. Environ*, 5 (3), 180-188.

-
- Lawless, H.T., & Heymann, H., (1998a). *Descriptive analysis, Sensory evaluation of food: principles and practices*, Chapman & Hall, pp. 341-378.
- Lawless, H.T., & Heymann, H., (1998b). *Introduction and overview, Sensory evaluation of food: principles and practices*, Chapman & Hall, pp. 1-27.
- Lawless, H.T., & Heymann, H., (1998c). *Sensory evaluation of food: principles and practices*. Chapman & Hall. New York.
- Lesschaeve, I., & Issanchou, S. (1996). Could selection tests detect the future performance of descriptive panellists? *Food Quality and Preference*, 7 (3-4), 177-183.
- Lu, R., & Abbott, J. (1997). Finite element modeling of transient responses of apples to impulse excitation. *Transactions of the ASAE*, 40 (5), 1395-1406.
- Lundahl, D.S., & McDaniel, M.R. (1988). The panelist effect - fixed or random? *Journal of Sensory Studies*, 3 (2), 113-121.
- Lundahl, D.S., & McDaniel, M.R. (1990). Use of contrasts for the evaluation of panel inconsistency. *Journal of Sensory Studies*, 5 (4), 265-277.
- M**acFie, H.J., Bratchell, N., Greenhoff, K., & Vallis, L.V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies*, 4 (2), 129-148.
- MacLeod, P., Sauvageot, F., & Köster, E.P., (2009). *Les caractéristiques d'une réponse sensorielles, Évaluation sensorielle: manuel méthodologique*, Lavoisier Tec & Doc, pp. 7-33.
- MacLeod, P., & Strigler, F., (2009). *Métrie sensorielle, Évaluation sensorielle: manuel méthodologique*, Lavoisier Tec & Doc, pp. 35-45.
- Mehinagic, E. (2004). *Recherche de descripteurs objectifs de la qualité sensorielle des pommes (Malus domestica)*. Université de Nantes - Faculté des Sciences et Techniques, Nantes.
- Mehinagic, E., Royer, G., Bertrand, D., Symoneaux, R., Laurens, F., & Jourjon, F. (2003). Relationship between Visible-NIR spectroscopy and sensory analysis of apples. *Food Quality and Preference*, 14, 473-484.
- Mehinagic, E., Royer, G., Symoneaux, R., Bertrand, D., & Jourjon, F. (2004). Prediction of the sensory quality of apples by physical measurements. *Postharvest biology and technology*, 34 (3), 257-269.
- Mehinagic, E., Royer, G., Symoneaux, R., & Jourjon, F. (2006). Relationship between apple sensory attributes and instrumental parameters of texture. *Journal of Fruit and Ornamental Plant Research*, 14 (Suppl. 2), 25-37.
- Meilgaard, M., Civille, G.V., & Carr, B.T., (1991a). 4 - *Factors influencing sensory verdicts, Sensory evaluation techniques*, CRC Press, pp. 37-42.
- Meilgaard, M., Civille, G.V., & Carr, B.T., (1991b). *Sensory evaluation techniques*. CRC Press.
-

- Molina-Delgado, D., Alegre, S., Barreiro, P., Valero, C., Ruiz-Altisent, M., & Recasens, I. (2009a). Addressing potential sources of variation in several non-destructive techniques for measuring firmness in apples. *Biosystems Engineering*, 104 (1), 33-46.
- Molina-Delgado, D., Alegre, S., Puy, J., & Recasens, I. (2009b). Relationship between acoustic firmness and Magness Taylor firmness in Royal Gala and Golden Smoothie apples. *Food Science and Technology International*, 15 (1), 31-40.
- Molnár, P.J. (1995). A model for overall description of food quality. *Food Quality and Preference*, 6 (3), 185-190.
- Monrozier, R., & Danzart, M. (2001). A quality measurement for sensory profile analysis The contribution of extended cross-validation and resampling techniques. *Food Quality and Preference*, 12 (5-7), 393-406.
- Moskowitz, H., (1983). *Product testing and sensory evaluation of foods* Marketing and R&D approaches. Food and Nutrition.Press Westport USA.
- Murray, J.M., Delahunty, C.M., & Baxter, I.A. (2001). Descriptive sensory analysis: past, present and future. *Food research international*, 34 (6), 461-471.
- N**æs, T. (1990). Handling individual differences between assessors in sensory profiling. *Food Quality and Preference*, 2 (3), 187-199.
- Næs, T. (1998). Detecting individual differences among assessors and differences among replicates in sensory profiling. *Food Quality and Preference*, 9 (3), 107-110.
- Næs, T., Brockhoff, P.B., & Tomic, O., (2010). *Statistics for sensory and consumer science*. John Wiley & Sons Ltd.Chichester.
- Næs, T., & Langsrud, Ø. (1998). Fixed or random assessors in sensory profiling? *Food Quality and Preference*, 9 (3), 145-152.
- O**raguzie, N., Alspach, P., Volz, R., Whitworth, C., Ranatunga, C., Weskett, R., & Harker, R. (2009). Postharvest assessment of fruit quality parameters in apple using both instruments and an expert panel. *Postharvest biology and technology*, 52 (3), 279-287.
- P**agès, J., & Husson, F., (2009). *Approche statistique de la performance en analyse sensorielle descriptive, Évaluation sensorielle: manuel méthodologique*, Lavoisier Tec & Doc, pp. 451-458.
- Pangborn, R. (1959). Influence of hunger on sweetness preferences and taste thresholds. *American Journal of Clinical Nutrition*, 7 (3), 280.
- Peirs, A., Lammertyn, J., Ooms, K., & Nicolai, B.M. (2001). Prediction of the optimal picking date of different apple cultivars by means of VIS/NIR-spectroscopy. *Postharvest biology and technology*, 21 (2), 189-199.
- Piggott, J.R. (1995). Design questions in sensory and consumer science. *Food Quality and Preference*, 6 (4), 217-220.

Pineau, N. (2006). *La performance en analyse sensorielle, une approche base de données.*, Université de Bourgogne, Dijon.

Pinheiro, J.C., & Bates, D.M., (2000). *Mixed-effects models in S and S-plus* ©. Springer Verlag. New York, LLC.

Plotto, A., McDaniel, M.R., & Mattheis, J.P. (1999). Characterization of Gala'Apple Aroma and Flavor: Differences between Controlled Atmosphere and Air Storage. *Journal of the American Society for Horticultural Science*, 124 (4), 416-423.

R Core Team, (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.R-project.org>.

Rizzolo, A., Vanoli, M., Spinelli, L., & Torricelli, A. (2010). Sensory characteristics, quality and optical properties measured by time-resolved reflectance spectroscopy in stored apples. *Postharvest biology and technology*, 58 (1), 1-12.

Romano, R., Brockhoff, P.B., Hersleth, M., Tomic, O., & Næs, T. (2008). Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Quality and Preference*, 19 (2), 197-209.

Rossi, F. (2001). Assessing sensory panelist performance using repeatability and reproducibility measures. *Food Quality and Preference*, 12 (5-7), 467-479.

Róth, E., Berna, A., Beullens, K., Yarramraju, S., Lammertyn, J., Schenk, A., & Nicolai, B. (2007). Postharvest quality of integrated and organically produced apple fruit. *Postharvest biology and technology*, 45 (1), 11-19.

Saftner, R.A., Abbott, J.A., Conway, W.S., Barden, C.L., & Vinyard, B.T. (2002). Instrumental and sensory quality characteristics of Gala'apples in response to prestorage heat, controlled atmosphere, and air storage. *Journal of the American Society for Horticultural Science*, 127 (6), 1006-1012.

Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin*, 2 (6), 110-114.

Sauvageot, F., (1982). *L'évaluation sensorielle des denrées alimentaires: aspects méthodologiques*. Lavoisier Tec & Doc.

Schlich, P. (1994). GRAPES: A Method and a SAS (R) Program for Graphical Representations of Assessor Performances. *Journal of Sensory Studies*, 9 (2), 157-170.

Schlich, P. (1997). CAP: Une méthode et un outil de contrôle rapide et synthétique des performances des sujets en évaluation sensorielle descriptive. In, *5èmes journées Agro-Industrie et Méthodes Statistiques*. INRA Versailles.

Schouten, R.E., Jongbloed, G., Tijssens, L.M.M., & van Kooten, O. (2004). Batch variability and cultivar keeping quality of cucumber. *Postharvest biology and technology*, 32 (3), 299-310.

Searle, S.R. (1995). An overview of variance component estimation. *Metrika*, 42 (1), 215-230.

- Seppä, L., Railio, J., Mononen, R., Tahvonen, R., & Tuorila, H. (2012). From profiles to practice: Communicating the sensory characteristics of apples to the wider audience through simplified descriptive profiles. *LWT - Food Science and Technology*, 47 (1), 46-55.
- Shmulevich, I., Galili, N., & Howarth, M. (2003). Nondestructive dynamic testing of apples for firmness evaluation. *Postharvest biology and technology*, 29 (3), 287-299.
- Sinesio, F., Moneta, E., & Saba, A. (1991). Comparison of multivariate methods of analysis to evaluate panellists' performance. *Food Quality and Preference*, 3, 201-201.
- Stevens, D.A. (1996). Individual differences in taste perception. *Food Chemistry*, 56 (3), 303-311.
- Stevens, M.A., & Albright, M. (1980). An approach to sensory evaluation of horticultural commodities [Fresh market tomatoes]. *HortScience*, 15 (1), 48-50.
- Stone, H., & Sidel, J. (1998). Sensory evaluation by quantitative descriptive analysis: development, applications, and the future. *IFT Annual Mtg*, 20-24.
- Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R.C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology*, 23-34.
- Stone, H., & Sidel, J.L., (2004). *Sensory evaluation practices*. Academic Press.
- Symoneaux, R., Royer, G., Madieta, E., Jourjon, F., Chollet, S., & Lombard, M. (2002). Differences Between the Two Opposite Sides of Bicoloured Apples: Sensory and Instrumental Measurements. In *International Conference Postharvest Unlimited*, ISHS, 383-387.
- Szczesniak, A.S. (2002). Texture is a sensory property. *Food Quality and Preference*, 13 (4), 215-225.
- Thybo, A.K., & Martens, M. (2000). Analysis of sensory assessors in texture profiling of potatoes by multivariate modelling. *Food Quality and Preference*, 11 (4), 283-288.
- Tijskens, L.M.M., Konopacki, P., & Simcic, M. (2003). Biological variance, burden or benefit? *Postharvest biology and technology*, 27 (1), 15-25.
- Tijskens, L.M.M., Schouten, R.E., Konopacki, P.J., Hribar, J., & Simčič, M. (2010). Modelling the biological variance of the yellow aspect of Granny Smith apple colour. *Journal of the Science of Food and Agriculture*, 90 (5), 798-805.
- Tomic, O., Nilsen, A., Martens, M., & Næs, T. (2007). Visualization of sensory profiling data for performance monitoring. *LWT-Food Science and Technology*, 40 (2), 262-269.
- Varela, P., Salvador, A., & Fiszman, S. (2005). Shelf-life estimation of 'Fuji' apples : Sensory characteristics and consumer acceptability. *Postharvest biology and technology*, 38 (1), 18-24.
- Varela, P., Salvador, A., & Fiszman, S. (2008). Shelf-life estimation of 'Fuji' apples : II. The behavior of recently harvested fruit during storage at ambient conditions. *Postharvest biology and technology*, 50 (1), 64-69.

Vaysse, P., & Landry, P., (2004). *Pomme-poire: de la récolte au conditionnement: outils pratiques*. Editions Centre technique interprofessionnel des fruits et légumes.

Vaysse, P., Masseron, A., & Trillot, M. (2000). Reconnaître les variétés de pommes et de poires/Recognizing apple and pear varieties. *Recherche*, 67, 02.

Wakeling, I.N., & MacFie, H.J.H. (1995). Designing consumer trials balanced for first and higher orders of carry-over effect when only a subset of k samples from t may be tested. *Food Quality and Preference*, 6 (4), 299-308.

Waldron, K.W., Smith, A.C., Parr, A.J., Ng, A., & Parker, M.L. (1997). New approaches to understanding and controlling cell separation in relation to fruit and vegetable texture. *Trends in Food Science & Technology*, 8 (7), 213-221.

Watada, A.E., & Abbot, J.A., (1985). *Apple Quality: Influences of Pre-and Postharvest Factors and Estimation by Objective Methods, Evaluation of Quality of Fruits and Vegetables*, AVI Publishing Co, pp. 63-81.

Watada, A.E., Abbott, J.A., & Hardenburg, R.E. (1980). Sensory characteristics of apple fruit. *Journal of the American Society for Horticultural Science*, 105 (3), 371-375.

Weibel, F., Bickel, R., Leuthold, S., & Alföldi, T. (1998). Are organically grown apples tastier and healthier? A comparative field study using conventional and alternative methods to measure fruit quality. In *Acta Hort*, ISHS, 417-426.

Williams, A.A., & Arnold, G.M. (1984). A new approach to the sensory analysis of foods and beverages. *Progress in Flavour Research*, 35-50.

Williams, A.A., & Carter, C.S. (1977). A language and procedure for the sensory assessment of Cox's Orange Pippin apples. *Journal of the Science of Food and Agriculture*, 28 (12), 1090-1104.

Williams, E.J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2 (2), 149-168.

Wolters, C.J., & Allchurch, E.M. (1994). Effect of training procedure on the performance of descriptive panels. *Food Quality and Preference*, 5 (3), 203-214.

Zude, M., Herold, B., Roger, J.M., Bellon-Maurel, V., & Landahl, S. (2006). Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. *Journal of food engineering*, 77 (2), 254-260.

Index des tableaux

Tableau 1. Principaux descripteurs de la texture des pommes	14
Tableau 2. Paramètres du modèle d'ANOVA standard	18
Tableau 3. Carrés moyens attendus selon la nature de l'effet sujet dans le modèle à deux facteurs produit et sujet et leur interaction	20
Tableau 4. Présentation des échantillons de pomme en évaluation sensorielle	39
Tableau 5: Caractéristiques des pommes utilisées	48
Tableau 6. Descripteurs sensoriels utilisés par le panel.....	50
Tableau 7. Performance inter-variétale – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des fruits dégustés	61
Tableau 8. Performance inter-variétale – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les sept descripteurs.....	63
Tableau 9. Performance inter-variétale – notes moyennes et erreurs standards pour les trois variétés de pommes et les sept descripteurs	64
Tableau 10. Performance inter-variétale – composantes de la variance et, entre parenthèses, leur pourcentage de la variation totale pour les sept descripteurs dans le cas du modèle (IV-1) aléatoire	64
Tableau 11. Performance intra-variétale – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des trois sous-lots dégustés pour chaque variété	67
Tableau 12. Performance intra-variétale – effet des facteurs variété, fermeté acoustique, sujet et des interactions variété*fermeté acoustique et sujet*variété*fermeté acoustique sur les notes des neuf produits (trois variétés de pomme*trois fermeté acoustique) et les sept descripteurs.....	72
Tableau 13. Performance intra-variétale - notes moyennes pour les trois variétés de pomme, les trois niveaux de fermeté acoustique et les sept descripteurs.....	72
Tableau 14. Test des effets des facteurs du modèle « fruit »	77
Tableau 15. Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs.....	81
Tableau 16. Composantes de la variance et, entre parenthèses, leur pourcentage de la variation totale pour les sept descripteurs dans le cas du modèle « fruit » (IV-11) aléatoire	81
Tableau 17. Comparaison des résultats d'ANOVA dans deux situations de variabilité biologique	85
Tableau 18. Démarche expérimentale pour l'observation des écarts de notation	89
Tableau 19. 2cart de notation – fermeté acoustique moyenne des lots initiaux et étendue de fermeté acoustique des fruits dégustés	89
Tableau 20. Effet du facteur sujet et de de l'interaction sujet*variété sur les tendances de notations pour les trois variétés de pomme et les sept descripteurs.....	91
Tableau 21. Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs – performance inter-variété	99
Tableau 22. Comparaison des résultats d'ANOVA pour l'effet de dispersion et le désaccord	104
Tableau 23. Coefficient de corrélation de Pearson entre les notes et les notes transformées pour chaque sujet et chaque répétition (n=120)	107
Tableau 24. Homogénéisation instrumentale – moyennes de fermeté mécanique (N) et écart-types des lots initiaux, hétérogènes et homogènes pour les trois variétés.....	112
Tableau 25. Homogénéisation instrumentale – moyennes de fermeté acoustique ($\text{Hz}^2\text{g}^{2/3}$) et écart-types des lots initiaux, hétérogènes et homogènes pour les trois variétés.....	112
Tableau 26. Écart-types de la réponse sensorielle pour les trois variétés, sept descripteurs et deux échantillonnages par pénétrométrie, et évolution des écart-types	115
Tableau 27. Coefficient de corrélation de Pearson entre les descripteurs sensoriels et la fermeté mécanique (n=18).....	115
Tableau 28. Échantillonnage hétérogène par pénétrométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs.....	116
Tableau 29. Échantillonnage homogène par pénétrométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs.....	116
Tableau 30. Écart-types de la réponse sensorielle pour les trois variétés de pomme, sept descripteurs et deux échantillonnages par sonométrie, et évolution des écart-types	118

Tableau 31. Coefficient de corrélation de Pearson entre les descripteurs sensoriels et les mesures instrumentales (n=18).....	118
Tableau 32. Étendues de fermeté acoustique pour les expérimentations de discrimination intra-variétale et d'homogénéisation par sonométrie.	119
Tableau 33. Échantillonnage hétérogène par sonométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs.....	121
Tableau 34. Échantillonnage homogène par sonométrie – Effet des facteurs variété, sujet, fruit et de l'interaction sujet*variété sur les notes des trois variétés de pomme et les sept descripteurs.....	121
Tableau 35. Écart-types de la réponse sensorielle pour les trois variétés de pommes, les cinq descripteurs et mode de présentation pour 1 et pour 4 sujets, et évolution des écart-types	123
Tableau 36. Écart-types de la réponse sensorielle pour les trois variétés de pommes, les cinq descripteurs et mode de présentation pour 1 et pour 8 sujets, et évolution des écart-types	123
Tableau 37. Écart-types de la réponse sensorielle pour les 5 descripteurs, 3 variétés et mode de présentation pour 4 et pour 8 sujets, et évolution des écart-types.....	124
Tableau 38. Mode '1 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs	125
Tableau 39. Mode '4 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs	125
Tableau 40. Mode '8 sujet par fruit' – effet des facteurs variété, sujet et de l'interaction sujet*variété sur la réponse sensorielle pour les trois variétés de pomme et les cinq descripteurs	125
Tableau 41. Composantes de la variance et , entre parenthèses, leur pourcentage de la variation totale avec méthode '4 sujets par fruit' et la méthode '8 sujets par fruit' pour les cinq descripteurs	128

Index des illustrations

Figure 1. Schéma général de la perception sensorielle.....	5
Figure 2. Étapes pour l'établissement d'un profil sensoriel (ISO, 1994)	13
Figure 3. Les différences d'utilisation de l'échelle entre les sujets	19
Figure 4. Le terme d'interaction : représentation graphique de ses deux interprétations.....	21
Figure 5. La variabilité biologique s'exprime à différents niveaux	25
Figure 6. Démarche expérimentale.....	47
Figure 7. Courbe force/déformation obtenue par pénétrométrie sur pomme non épluchée en utilisant la machine de traction universelle (Mehinagic, 2004)	52
Figure 8. Sonomètre (AFS; AWETA, Nootdorp, The Netherlands) (à gauche) et schéma du dispositif (à droite). 53	
Figure 9. Emplacement des mesures acoustiques	53
Figure 10. Caractérisation des fruits et répartition en cinq groups (A à E) – exemple de la variété Ariane	60
Figure 11. Performance inter-variétale – distribution de l'indice acoustique de fermeté pour les trois variétés 60	
Figure 12. Performance inter-variétale –estimations des composantes de la variance pour les sept descripteurs	65
Figure 13. Performance intra-variétale – distribution de l'indice acoustique de fermeté pour les trois variétés 67	
Figure 14. Notes moyennes pour le descripteur <i>croquant</i> , les trois variétés et les trois niveaux de fermeté acoustique.....	70
Figure 15. Décomposition de la variance selon le modèle standard et le modèle « fruit » (d'après Pineau (2006))	78
Figure 16. Estimations des composantes de la variance pour le modèle « fruit » (IV-11) et les sept descripteurs	82
Figure 17. Passage du modèle standard au modèle « fruit » - discrimination.....	83
Figure 18. Passage du modèle standard au modèle « fruit » - répétabilité	83
Figure 19. Passage du modèle standard au modèle « fruit » - désaccord total.....	83
Figure 20. Passage du modèle standard au modèle « fruit » - effet niveau	83
Figure 21. Distribution des données simulées pour des variabilités intra-lot faible (à gauche) et forte (à droite)	84
Figure 22. Statistiques de test pour les données simulées à variabilité intra-lot faible pour le modèle standard et le modèle fruit.....	86
Figure 23. Statistiques de test pour les données simulées à variabilité intra-lot forte pour le modèle standard et le modèle fruit.....	86
Figure 24. Différence à la moyenne les trois variétés de pomme et pour le descripteur <i>croquant</i>	92
Figure 25. Évaluation moyenne individuelle des trois variétés pour les deux sous-panels et les trois répétitions pour le descripteur <i>croquant</i>	94
Figure 26. Décomposition de la variance selon le modèle « fruit » et l'approche « assessor » du modèle « fruit »	97
Figure 27. Passage du modèle « fruit » à l'approche « assessor » - désaccord total.....	100
Figure 28. Passage du modèle « fruit » à l'approche « assessor » - discrimination.....	100
Figure 29. Passage du modèle « fruit » à l'approche « assessor » - répétabilité.....	100
Figure 30. Passage du modèle « fruit » à l'approche « assessor » - effet de niveau	100
Figure 31. Passage du modèle « fruit » à l'approche « assessor » - effet fruit	101
Figure 32. Estimations des composantes de la variance pour l'approche « assessor » du modèle « fruit » (IV-18) et les sept descripteurs	101
Figure 33. Scénario 1 avec dispersion et sans désaccord - Moyenne pour chaque sujet et chaque variété – les sujets 11 à 15 utilisent une plus large étendue de l'échelle ; les sujets 16 à 20 utilisent une plus faible étendue d'échelle	102
Figure 34. Scénario 2 avec désaccord et dispersion - Moyenne pour chaque sujet et chaque variété – les sujets 12, 14, 16, 18 et 20 présentent un classement des produits inverse au classement moyen ($A = 4,4 < B = 4,9 < C = 5,7$)	103
Figure 35. Statistiques de test pour les données simulées dans le cas du scénario 1 pour le modèle standard « fruit » et l'approche « assessor » du modèle « fruit »	105

Figure 36. Statistiques de test pour les données simulées dans le cas du scénario 2 pour le modèle standard « fruit » et l'approche « assessor » du modèle « fruit » 105

Figure 37. Estimation des composantes de la variance pour le modèle « fruit » et l'approche « assessor » du modèle « fruit » pour les données simulées scénario 1 (à gauche) et scénario 2 (à droite) 106

Figure 38. Méthodes d'échantillonnage hétérogène et homogène 111

Figure 39. Note individuelles des sujets en fonction des notes moyennes des fruits pour la variété Ariane et descripteur fermé pour le lot hétérogène (à gauche) et le lot homogène (à droite) 119

Annexes

Annexe 1. Données simulées pour évaluer le modèle « fruit »

Variété	Fruit	Sujet	Variabilité biologique faible	Variabilité biologique forte
A	1	1	3,3	2.0
A	1	2	4,9	2.8
A	2	3	3,4	5.0
A	2	4	5,1	6.2
A	3	5	3,7	9.8
A	3	6	5,7	9.7
A	4	7	4,5	3.5
A	4	8	4,7	4.4
A	5	9	3,6	7.4
A	5	10	5,2	8.4
A	6	11	2,8	3.0
A	6	12	3,7	5.0
A	7	13	4,7	3.1
A	7	14	5,6	3.8
A	8	15	4,9	4.9
A	8	16	6,5	6.8
A	9	17	3,4	6.6
A	9	18	5,0	7.7
A	10	19	4,1	4.0
A	10	20	4,7	4.3
A	11	1	4,1	2.0
A	11	2	4,2	2.0
A	12	3	3,8	1.7
A	12	4	4,1	2.6
A	13	5	3,7	3.0
A	13	6	5,6	3.1
A	14	7	4,0	3.3
A	14	8	4,5	3.4
A	15	9	4,4	1.3
A	15	10	5,0	2.3
A	16	11	3,9	4.6
A	16	12	4,6	6.0
A	17	13	3,1	0.3
A	17	14	4,6	0.3
A	18	15	5,6	3.0
A	18	16	5,9	4.0
A	19	17	4,4	7.0
A	19	18	5,2	7.3
A	20	19	4,7	6.9
A	20	20	6,3	7.7

B	21	1	4,5	8.3
B	21	2	5,8	9.2
B	22	3	7,1	8.3
B	22	4	7,9	9.5
B	23	5	6,1	2.6
B	23	6	8,0	3.3
B	24	7	5,6	3.3
B	24	8	6,8	4.0
B	25	9	4,8	5.6
B	25	10	5,2	5.7
B	26	11	6,2	6.4
B	26	12	7,3	7.9
B	27	13	5,8	8.2
B	27	14	7,5	10.0
B	28	15	4,5	5.6
B	28	16	5,3	7.0
B	29	17	5,5	3.4
B	29	18	6,3	4.2
B	30	19	6,2	9.0
B	30	20	7,3	9.7
B	31	1	5,7	1.6
B	31	2	6,2	2.4
B	32	3	4,9	7.7
B	32	4	5,1	7.7
B	33	5	5,3	5.8
B	33	6	5,7	7.6
B	34	7	5,1	4.3
B	34	8	6,8	5.6
B	35	9	4,3	1.3
B	35	10	6,1	1.4
B	36	11	6,0	6.8
B	36	12	6,7	8.2
B	37	13	5,5	9.7
B	37	14	6,3	9.7
B	38	15	4,8	2.9
B	38	16	6,7	4.1
B	39	17	6,3	3.3
B	39	18	6,9	4.5
B	40	19	5,4	6.1
B	40	20	6,0	8.0

Annexe 2. Données simulées pour évaluer l'approche « assessor »

Variété	Fruit	Sujet	Normal	Dispersion	Disagree	
A		1	1	1,5	1,6	1,7
A		1	2	2,7	2,5	2,6
A		2	3	2,8	2,7	2,5
A		2	4	3,3	3,7	3,5
A		3	5	1,5	1,3	1,4
A		3	6	2,5	2,3	2,3
A		4	7	2,0	2,0	1,8
A		4	8	2,6	2,8	2,9
A		5	9	3,2	3,0	2,9
A		5	10	4,0	4,1	3,8
A		6	11	4,2	3,8	4,1
A		6	12	5,4	4,6	7,7
A		7	13	6,3	5,7	6,4
A		7	14	7,2	6,6	9,9
A		8	15	1,4	0,9	1,5
A		8	16	2,2	2,8	4,8
A		9	17	3,6	3,9	3,7
A		9	18	4,3	5,0	6,7
A		10	19	2,8	3,3	2,5
A		10	20	3,8	4,1	6,1
A		11	1	3,0	2,9	2,7
A		11	2	3,4	3,8	3,8
A		12	3	3,7	3,5	3,6
A		12	4	4,5	4,7	4,4
A		13	5	2,6	2,2	2,4
A		13	6	3,5	3,2	3,3
A		14	7	2,7	2,8	3,0
A		14	8	3,7	3,9	3,7
A		15	9	3,9	4,2	4,0
A		15	10	4,9	5,0	4,8
A		16	11	5,4	5,0	5,3
A		16	12	6,3	5,6	8,6
A		17	13	7,6	6,7	7,3
A		17	14	8,2	7,8	11,0
A		18	15	2,4	1,8	2,3
A		18	16	3,4	3,9	5,8
A		19	17	4,3	5,2	4,5
A		19	18	5,5	5,7	7,9
A		20	19	3,5	4,0	3,5
A		20	20	4,6	4,9	7,1
B		21	1	2,8	2,8	3,0
B		21	2	3,6	3,7	3,9
B		22	3	3,7	3,5	3,8
B		22	4	4,6	4,3	4,3

B	23	5	2,6	2,2	2,4
B	23	6	3,1	3,4	3,3
B	24	7	2,7	2,7	3,0
B	24	8	4,0	3,9	3,9
B	25	9	4,2	4,1	4,2
B	25	10	4,7	5,1	5,1
B	26	11	5,1	5,2	5,5
B	26	12	6,4	6,3	6,1
B	27	13	7,6	7,4	7,3
B	27	14	8,3	8,2	8,1
B	28	15	2,5	2,3	2,5
B	28	16	3,4	3,2	3,5
B	29	17	4,7	4,6	4,7
B	29	18	5,3	5,5	5,4
B	30	19	3,9	3,6	3,8
B	30	20	4,6	4,9	4,7
B	31	1	3,7	3,9	4,0
B	31	2	4,7	4,9	4,5
B	32	3	4,6	4,4	4,5
B	32	4	5,4	5,7	5,3
B	33	5	3,5	3,5	3,1
B	33	6	4,2	4,4	4,2
B	34	7	3,9	4,1	4,1
B	34	8	4,6	4,7	5,0
B	35	9	5,2	5,1	5,0
B	35	10	5,8	6,0	6,2
B	36	11	6,2	6,4	6,3
B	36	12	7,4	7,0	7,2
B	37	13	8,1	8,5	8,4
B	37	14	9,1	9,2	9,2
B	38	15	3,7	3,3	3,5
B	38	16	4,2	4,5	4,2
B	39	17	5,7	5,7	5,3
B	39	18	6,2	6,4	6,5
B	40	19	4,6	4,6	4,6
B	40	20	5,5	5,6	5,5
C	41	1	4,3	4,4	4,1
C	41	2	5,0	5,1	5,0
C	42	3	4,9	5,3	4,9
C	42	4	5,9	6,2	5,7
C	43	5	4,0	4,0	4,0
C	43	6	4,7	4,9	5,0
C	44	7	4,6	4,2	4,4
C	44	8	5,3	5,2	5,3
C	45	9	5,3	5,4	5,7
C	45	10	6,3	6,5	6,4

C	46	11	6,6	7,4	6,8
C	46	12	7,8	8,0	5,3
C	47	13	8,6	9,5	9,0
C	47	14	9,6	10,4	7,2
C	48	15	4,2	4,3	4,1
C	48	16	5,0	4,4	2,4
C	49	17	6,2	5,4	6,1
C	49	18	7,0	6,5	4,4
C	50	19	5,0	4,8	5,4
C	50	20	6,1	5,8	3,6
C	51	1	5,4	5,4	5,1
C	51	2	6,0	6,0	6,2
C	52	3	6,1	6,1	6,1
C	52	4	7,0	6,8	6,7
C	53	5	5,0	4,7	4,9
C	53	6	6,0	6,0	5,7
C	54	7	5,6	5,2	5,3
C	54	8	6,2	6,1	6,4
C	55	9	6,4	6,7	6,4
C	55	10	7,6	7,3	7,4
C	56	11	8,0	8,5	8,0
C	56	12	8,5	9,1	6,0
C	57	13	9,6	10,4	10,0
C	57	14	11,0	11,3	8,3
C	58	15	5,1	5,6	5,1
C	58	16	5,9	5,1	3,2
C	59	17	7,2	6,7	6,9
C	59	18	7,7	7,3	5,3
C	60	19	6,0	5,7	6,0
C	60	20	6,9	6,5	4,7

Annexe 3. Publication et Communications par poster

Bavay, C., Symoneaux, R., Maître, I., Kuznestova, A., Brockhoff, P.B., Mehinagic, E. (2013). Adaptation of sensory evaluation methodology for variable products: a study on apple. *Postharvest biology and technology*, 77, 67-74.

Bavay, C., Symoneaux, R., Maître, I., Courcoux, P., Mehinagic, E. (2011). Non-destructive sampling to improve sensory descriptive results of varying products: an application on apple. In *9th Pangborn Sensory Science Symposium*, 4-8 September, Toronto, Canada

Bavay, C., Symoneaux, R., Maître, I., Mehinagic, E. (2012). Management of fruit presentation in sensory evaluation of apples for more reliable results. In *II International Symposium on Horticulture in Europe*, 1-5 July, Angers, France

Bavay, C., Symoneaux, R., Maître, I., Mehinagic, E. (2012). Sensory profile adaptation to biological variability of fruit batch. In *5th European Conference on Sensory and Consumer Research*, 9-12 September, Bern, Switzerland

Bavay, C., Symoneaux, R., Maître, I., Kuznestova, A., Brockhoff, P.B., Mehinagic, E. (2012). The importance of methodological adaptation to get reliable results in sensory evaluation of product presenting biological variability: a case study on apple. In *SSP conference*, 10-12 October, Jersey City NJ, USA



Importance of fruit variability in the assessment of apple quality by sensory evaluation

Cécile Bavay^{a,*}, Ronan Symoneaux^a, Isabelle Maître^a, Alexandra Kuznetsova^b,
Per Bruun Brockhoff^b, Emira Mehinagic^a

^a LUNAM Université, SFR QUASAV 4207, Groupe ESA, UPSP GRAPPE, 55, rue Rabelais BP 30748, F-49007 Angers Cedex 01, France

^b DTU Informatics, Statistical Section, Technical University of Denmark, Richard Petersens Plads, Building 305, DK-2800 Kongens Lyngby, Denmark

ARTICLE INFO

Article history:

Received 13 June 2012

Accepted 13 November 2012

Keywords:

Apple
Mixed hierarchical model
Panel performance
Sensory evaluation
Variability

ABSTRACT

The assessment of produce quality is a major aspect of applied postharvest biology. Horticultural researchers working on organoleptic quality of fruit need objective methods for the evaluation of sensory properties. The development of sensory methodologies specifically for apples highlighted the problem of handling variation due to fruit variability and assessor differences. The aim of this study was to investigate the weight of within-batch variability in sensory evaluation of apples and to propose a methodology that accounts for this variability. Prior to sensory analysis, for three apple cultivars, apples were sorted into homogenous acoustic firmness categories within each cultivar. The discrimination ability of the trained panel was observed not only between cultivars but also within each cultivar for crunchiness, firmness, juiciness and acidity. Following these results, a mixed hierarchical model for the analysis of the sensory data was proposed to measure the contribution of fruit variability to the variability of sensory scores. The results showed the efficiency of the model in quantifying within-batch variability. Fruit sampling and presentation methods as well as data handling procedures are suggested for obtaining reliable sensory results in the assessment of apple quality.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The assessment of produce quality is a major aspect of applied postharvest biology. Selection of new apple cultivars as well as pre- and postharvest quality management needs objective methods for the evaluation of the organoleptic properties. Indeed, one of the most important aspects of apple quality for consumers is the sensory quality (Harker et al., 2003), more specifically texture and taste (Dailliant-Spinnler et al., 1996; Jaeger et al., 1998). These sensory qualities of apple are conventionally assessed by trained panels (Meilgaard et al., 1991; Lawless and Heymann, 1998).

However, sensory assessments include an important variability in their results. One cause of this variability comes from assessor differences; there are inherent variations between assessors such as genetics and use of the scale. This issue has been raised by many authors and investigated by postharvest scientists (Oraguzie et al., 2009; Brookfield et al., 2011) as well as statisticians (Brockhoff and Skovgaard, 1994; Næs, 1998; Brockhoff, 2003; Romano et al., 2008). Traditionally, assessor differences are observed through the

use of a two-way analysis of variance (ANOVA) model with sample and assessor main effects together with the sample-by-assessor interaction.

In addition to this, Mehinagic et al. (2003) suggested that assessor differences could be explained by apple heterogeneity because each assessor tasted different fruit. This within-batch variability has been pointed out as a source of variability in results from the early development of sensory evaluation for horticultural products (Williams and Carter, 1977; Stevens and Albright, 1980; Heintz and Kader, 1983). As emphasized by Hampson et al. (2000), differences within a sample may make differences between samples more difficult to detect. High variability within each batch and even within each fruit creates a demand for more reliable methods.

The within-batch variability is often treated as an *ever-present* problem which should be minimised as much as possible. The most commonly used technique for the reduction of within-batch variability is prior sorting and grading on external quality attributes (Tijsskens et al., 2003) such as size, shape, colour and absence of damage or blemish (Hampson et al., 2000; King et al., 2000; Konopacka and Plocharski, 2004; Ioannides et al., 2007; Montero-Prado et al., 2011). Instrumental measurements including a puncture test, soluble solids content, Acoustic Firmness Sensor and Sinclair Internal Quality Firmness Tester have also been used to homogenise batches (Harker et al., 2002a, 2008; Konopacka and Plocharski, 2004).

* Corresponding author at: GROUPE ESA – Unité de Recherche GRAPPE, Ecole Supérieure d'Agriculture d'Angers, 55 Rue Rabelais – BP 30748, 49007 ANGERS Cedex 01, France. Tel.: +33 2 41 23 55 55; fax: +33 2 41 23 55 00.

E-mail address: c.bavay@groupe-esa.com (C. Bavay).

In addition, homogenization attempts have been done by sharing a tested apple among several assessors or by using small parts of the same fruit over several treatments. Oraguzie et al. (2009) and Brookfield et al. (2011) are examples of the first case. They used the sharing of apples to compare assessor assessment. Abbott et al. (2004) and Bhagwat et al. (2004) homogenized samples of treated fresh-cut apple for each assessor. In these two studies, each assessor evaluated samples of untreated and treated slices from the same apple. Thus, for each assessor, within-batch variability was limited. The efficiency of the aforementioned methods to reduce within-batch variability has not been proved nor quantified.

In order to quantify within-batch variability, statistical tools are needed. Mixed effect models present a statistical framework that allows a simultaneous characterization of the main effects that influence an experiment, together with the estimation of the different components that affect the intrinsic total score variability (Pinheiro and Bates, 2000). Moreover, within-batch variation has become a major topic in postharvest technology (Hertog, 2002; Tijksens et al., 2003; Hertog et al., 2004; Schouten et al., 2004; De Ketelaere et al., 2006; Aguirre et al., 2008; Bobelyn et al., 2010), however, this has not yet been applied in sensory evaluation.

The main issue of the present work was to propose a methodology, for postharvest researchers using sensory tools, to account for the within-batch variability in sensory evaluation in order to get more reliable conclusions. The focus of this paper is a specific design for sensory experiments and the corresponding data analysis, that allow for quantification of this within-batch variability. To fulfil our objectives, the sensory description of three apple cultivars was investigated according to seven sensory attributes of texture and taste. The study first considered the ability of a trained panel to sense small differences between apples within a single batch of a given cultivar (previously categorized by an instrumental measurement) to then adapt a model for variance analysis that accounts for within-batch variability. Contributions of this resulting model compared to the standard model were investigated.

2. Materials and methods

2.1. Fruit

Three different apple (*Malus domestica*) cultivars, namely 'Ariane', 'Braeburn' and 'Pink Lady®', were studied. Six hundred fruit of each cultivar were harvested between September and October 2010 at commercial maturity. The fruit were provided by a local wholesaler and stored at 4 °C, and were brought to room temperature 24 h before tasting.

2.2. A priori fruit categorization

In order to categorize apples within each cultivar for intra-cultivar discrimination evaluation, an instrumental measurement was needed. Even if Magness–Taylor firmness measurement has been widely studied and significantly related to sensory quality, the partial destruction of the samples constituted a large disadvantage for apple-to-apple evaluation. Therefore, a non-destructive categorization was needed.

Non-destructive instruments, such as the Acoustical Firmness Sensor (AFS) and Sinclair Internal Quality Firmness Tester (SIQ-FT) were used by Harker et al. (2008) to categorize apples within single cultivars according to texture properties. The applied categorizations were found to be effective in segregating the apples and the consumer responses obtained were as wide as when studying different cultivars. Moreover, the acoustic firmness measured by the AFS represents the mechanical properties of the entire fruit, unlike the firmness measured by the SIQ-FT which is localized. In addition,

acoustic firmness has been used to measure changes during storage (De Belie et al., 2000) as well as maturity differences (Mehinagic et al., 2006; Molina-Delgado et al., 2009a), and it has been significantly related to sensory properties (Abbott et al., 1992; Abbott and Lu, 1996; Barreiro et al., 1998; Symoneaux et al., 2005). For these reasons, the Acoustic Firmness Sensor (AFS; AWETA, Nootdorp, The Netherlands) was chosen as a truly non-destructive method allowing for rapid measurement of the overall texture of the apple.

With this device, an acoustic signal is generated by means of gentle impact on the fruit. This signal is then processed and transformed to obtain a peak of natural frequency, which is used to calculate the acoustic firmness (S), also called the stiffness index, from the following equation (Chen and De Baerdemaeker, 1993):

$$S = f^2 \times m^{2/3} \quad (1)$$

where f is the first resonance frequency (Hz) and m is the fruit mass (g).

Fruit acoustic firmness was measured at the equator in four replicates. Each replicate was measured after the apple was rotated by 90° (Chen and DeBaerdemaeker, 1993). The four replicates were averaged to obtain one mean value of acoustic firmness per apple. Then apples were sorted and classified according to their mean acoustic firmness. For each cultivar, the total range of acoustic firmness (R) was divided into five classes of equal range ($R/5$). Thus, five groups of apples were determined (A through E in an increasing order of acoustic firmness values). In order to favour perceptible sensory differences between the tested apples, only groups A, C and E were tasted.

2.3. Sensory evaluation

The sensory panel included 19 assessors trained at the research unit GRAPPE (Groupe ESA, Angers). These assessors were recruited from April to May 2010 and trained until October 2010 at the rate of two 1.5 h sessions per week. The total training was completed in 37.5 h. Selection and training were conducted according to the recommendations of ISO standards (ISO, 1994). The sensory attributes studied were *crunchiness*, *firmness*, *crispness*, *juiciness*, *fondant*, *acidity* and *sweetness*, as defined in Table 1. These chosen attributes had been previously identified as main indicators of apple quality (Daillant-Spinnler et al., 1996; Jaeger et al., 1998; Symoneaux et al., 2012). After the sensory attributes were well understood, the assessors were trained to detect small differences by paired comparison (ISO, 2005) and sensory profiling with feedback (Findlay et al., 2007). For the evaluations, apples were cut into eight slices. Each assessor received two slices of an apple so that the same fruit was distributed to four assessors to take into account the within-batch variability. The washed unpeeled apple slices were randomly presented to the assessors, under red-light illumination and at room temperature. Assessors were asked to bite the slice without eating the peel. A continuous non-structured scale was used for evaluation; the left end corresponding to the lowest intensity (value 0) and the right end to the highest intensity (value 10). Each assessor rinsed his/her mouth with mineral water and ate unsalted crackers between sample analyses.

2.4. Experimental design

Three experiments were designed to gradually increase the level of discrimination:

- I. Three samples (four replicates each) of the three different cultivars, each with a medium acoustic firmness level, were presented to the assessors (19 assessors from the panel participated in this experiment);

Table 1
List of the sensory attributes of apples with definitions.

Attribute	Definition	References
Crunchiness	Noise produced at the first bite by the incisors pushed into the flesh and the breakdown of the apple piece	Low: Golden Delicious placed at room temperature for 2 weeks – High: Pink Lady® stored at 4 °C
Firmness	Resistance of the flesh to molars pressure during the first chew	Low: Elstar stored at 4 °C – High: Granny Smith stored at 4 °C
Crispness	Noise produced by the gridding of the apple piece during chews 2 to 5	Low: Golden Delicious placed at room temperature for 2 weeks – High: Tentation stored at 4 °C
Juiciness	Amount of liquid released during chews 2 to 5	Low: Golden Delicious stored at 4 °C – High: Pink Lady® stored at 4 °C
Fondant	Capacity of apple pieces to dissolve after 5 chews	Low: Golden Delicious stored at 4 °C – High: Golden Delicious placed at room temperature for 2 weeks
Acidity	One of the basic tastes (e.g., malic acid)	Low: 1.3 g/L malic acid and 68 g/L saccharose – High: 8.0 g/L malic acid and 68 g/L saccharose
Sweetness	One of the basic tastes (e.g., saccharose)	Low: 68 g/L saccharose and 1.3 g/L malic acid – High: 140 g/L saccharose and 1.3 g/L malic acid

Table 2
Range of acoustic firmness (Hz² g^{2/3}) for the different samples presented to the panel.

Experiment	Acoustic firmness level	Ariane	Braeburn	Pink Lady®
I.	Medium	[27.7; 31.1]	[30.8; 32.3]	[31.2; 32.5]
II.	Low	[14.0; 15.9]	[18.3; 21.8]	[22.9; 26.8]
II.	High	[29.5; 31.6]	[33.8; 35.5]	[31.8; 35.6]
III.	Low	[9.6; 11.5]	[15.1; 18.9]	[22.1; 22.8]
III.	Medium	[17.3; 23.0]	[25.3; 26.1]	[27.4; 28.1]
III.	High	[28.6; 29.5]	[33.6; 35.4]	[31.7; 32.0]

- II. Six samples (four replicates each) of the three cultivars, each with low and high acoustic firmness levels, were presented to the assessors (19 assessors from the panel participated in this experiment);
- III. Nine samples (two replicates each) of the three cultivars, each with low, medium and high acoustic firmness levels, were presented to the assessors (18 assessors from the panel participated in this experiment).

Experiment I was designed to measure the assessors' capacity to discriminate between the apple cultivars. Experiments II and III were designed to verify if the assessors were able to further discriminate and sense differences in apples of the same cultivar with different acoustic firmness levels. Randomization was done across all cultivars and acoustic firmness levels.

The acoustic firmness ranges for each cultivar and each experiment are presented in Table 2. For each cultivar, the average acoustic firmness values differed for the three experiments because they were conducted from November to December 2010.

2.5. Statistical analysis

The sensory data were analysed by analysis of variance (ANOVA). Details of these analyses are provided below. All the analyses were performed with the R software package (version 2.14.2) (R Core Team). Since the models included random effects, they were

fitted with the lmer function of the lme4 package, and p values were calculated with the step function of the lmerTest package (Kuznetsova et al., 2012). For the fixed effect, p-value calculation was based on the F test using Satterthwaite's approximation; for the random effects, p-values were calculated based on the Log likelihood ratio test.

For each analysis, a significance level of 5% was considered. Effects with significance levels that ranged from 5% to 10% were identified as trends.

2.5.1. Panel performances

First, to demonstrate the panel's discriminatory ability, their performance for each attribute was observed (as a whole panel). In sensory profiling, performance is usually checked by doing a profiling task and testing for discrimination, scoring level and agreement between assessors. For random replicates with no systematic structure among them, a two-way mixed model that includes a fixed sample effect, a random assessor effect and interaction is advised (Næs et al., 2010).

For experiment I, differences would be mainly due to the different cultivars. Here, a mixed ANOVA model was used:

$$X_{ijk} = \alpha_i + B_j + C_{ij} + \varepsilon_{ijk} \quad (2)$$

where $\varepsilon_{ijk} \sim N(0, \sigma^2)$, $B_j \sim N(0, \sigma_{\text{assessor}}^2)$ and $C_{ij} \sim N(0, \sigma_{\text{cultivar} \times \text{assessor}}^2)$; all terms are independent.

The fixed cultivar effect α_i represents the differences between the average scores of the different cultivars. The random assessor effect B_j represents the differences in the scoring levels between assessors. The interaction between the cultivar and the assessor C_{ij} represents the assessors' disagreement and differences in the range of the scores of the cultivars. Post hoc analyses were conducted to observe differences between the cultivars.

For experiments II and III, sample differences would be given by different cultivars with different acoustic firmness levels. A mixed ANOVA model was also applied:

$$X_{ijkl} = \alpha_i + B_j + \delta_l + \nu_{jl} + C_{ijl} + \varepsilon_{ijkl} \quad (3)$$

where $\varepsilon_{ijkl} \sim N(0, \sigma^2)$, $B_j \sim N(0, \sigma_{\text{assessor}}^2)$ and $C_{ijl} \sim N(0, \sigma_{\text{cultivar} \times \text{acoustic firmness} \times \text{assessor}}^2)$; all terms are independent.

The fixed cultivar effect α_i represents the differences between the average scores of the different cultivars. The fixed acoustic firmness δ_l represents the differences between the average scores for the different acoustic firmness levels. The fixed interaction between the cultivar and acoustic firmness ν_{jl} represents differences in the ranking of the acoustic firmness levels for the three cultivars. The random assessor effect B_j represented differences in the scoring level between assessors. The interaction between cultivar, acoustic firmness, and assessor C_{ijl} represents the assessors' disagreement and differences in the range of the scores. Post hoc analyses were conducted to observe the differences between cultivars as well as between acoustic firmness levels within each cultivar.

2.5.2. Including fruit variability in the model

Following the confirmation of the panel's discriminatory ability within a cultivar (results presented in Sections 3.2 and 3.3), accounting the within-batch variability in the total scores variability was important. Thus, adaptation of the statistical analyses of sensory scores was needed. For this purpose, data from experiment III were used since it included the three cultivars with a large within-batch variability. The acoustic firmness levels were no longer considered as a studied factor. The data set was therefore the result from the sensory description of three cultivars tasted by 18 assessors in six replicates. Replicates were the result of grouping the three firmness levels evaluated in two replicates. Considering

Table 3
Significance levels for the sensory evaluation of the three apples cultivars.

Attribute	Cultivar	Assessor	Assessor × cultivar
Crunchiness	<0.001	<0.001	1.000
Firmness	<0.001	<0.001	1.000
Crispness	<0.001	<0.001	0.353
Juiciness	<0.001	0.006	0.209
Fondant	<0.001	<0.001	1.000
Acidity	<0.001	<0.001	0.992
Sweetness	0.076	<0.001	0.587

Table 4
Mean scores of the seven sensory attributes for the three apple cultivars.

Attributes	Ariane	Braeburn	Pink Lady®
Crunchiness	5.45a	6.00b	8.02c
Firmness	4.22a	4.49a	7.51b
Crispness	4.75a	5.75b	7.45c
Juiciness	5.07a	5.93b	7.33c
Fondant	5.06b	4.41b	1.63a
Acidity	5.46c	3.78a	3.95b
Sweetness	4.96b	4.33a	4.57ab

Within each row, values followed by the same letters are not significantly different ($\alpha=0.05$).

the large variability of each cultivar, we proposed to take it into account introducing a random fruit effect, as shown in Eq. (4):

$$X_{ijkl} = \alpha_i + B_j + C_{ij} + D_{l(ijk)} + \varepsilon_{ijkl} \quad (4)$$

where $B_j \sim N(0, \sigma_{\text{assessor}}^2)$, $C_{ij} \sim N(0, \sigma_{\text{cultivar} \times \text{assessor}}^2)$, $D_{l(ijk)} \sim N(0, \sigma_{\text{fruit}}^2)$, $\varepsilon_{ijkl} \sim N(0, \sigma^2)$; all terms are independent.

The results of this model were compared to the results from the standard model without the inclusion of a fruit effect:

$$X_{ijk} = \alpha_i + B_j + C_{ij} + \varepsilon_{ijk} \quad (5)$$

where $B_j \sim N(0, \sigma_{\text{assessor}}^2)$, $C_{ij} \sim N(0, \sigma_{\text{cultivar} \times \text{assessor}}^2)$, $\varepsilon_{ijk} \sim N(0, \sigma^2)$; all terms are independent.

The inclusion of the fruit effect was made possible by having several assessors tasting the same apple.

3. Results and discussion

The panel performance and more precisely, their discriminatory ability, was studied in three different experiments. The aim was to demonstrate the ability of a sensory panel to discriminate between cultivars and within each cultivar batch. For this purpose ANOVAs were first performed independently for each of the seven sensory attributes studied. Secondly, taking into account these results, a model accounting for within-batch variability was applied and results were compared to those from the currently standard model. All results are presented and discussed in the subsections below.

3.1. Discrimination between cultivars: exp I

As expected, the panel was able to sense differences between the three cultivars. The results from ANOVA (Table 3) showed the effects of the cultivar to be significant for all sensory attributes except for *sweetness*, which still showed a trend. The mean scores for each cultivar are presented in Table 4. The three cultivars were distinguished well for *crunchiness*, *crispness*, *juiciness* and *acidity*. The 'Ariane' and 'Braeburn' apples showed similarities in *firmness* and *fondant*. With regard to taste, 'Ariane' apples were the most acidic and the sweetest, whereas 'Braeburn' apples were the least acidic and the least sweet. These results are in line with the cultivar description given by Vaysse and Landry (2004). A significant assessor effect was observed for all sensory attributes. This effect

reveals a level effect with regard to difference between assessors in their average scoring. The results also showed non-significant cultivar-by-assessor interactions, which illustrated agreement on cultivar ranking by assessors.

3.2. Discrimination between cultivars, each with low and high acoustic firmness: exp II

In addition to being able to differentiate between cultivars, the panel was able to detect differences between apples with low and high acoustic firmness. The ANOVA results showed that the cultivars have different characteristics for all attributes and that the assessors were able to discriminate between fruit with different acoustic firmness levels for *crunchiness*, *firmness*, *crispness*, *juiciness* and *acidity* (Table 5). Trends of discrimination were observed for *fondant* and *sweetness*. Within-batch variation can be described as the composite of biological properties that differentiate individual units of a batch (adapted from Tijskens and Konopacki (2003)). One of those properties could be maturity (Kingston, 1991). Results of discrimination within cultivars would thus be in line with results demonstrating the ability of a sensory panel to distinguish between apples with different maturity levels (Watada et al., 1980; Harker et al., 2002b; Seppä et al., 2012).

The interaction between cultivar and acoustic firmness was significant for *crunchiness*, *firmness*, *crispness* and *juiciness*, and showed a trend for *fondant* and *acidity*, which indicates that differences between acoustic firmness levels depended on the cultivar. The mean scores, presented in Table 6, confirmed that the discrimination depends on the cultivar. The *crunchiness*, *firmness*, *crispness* and *fondant* for the low and high acoustic firmness levels differed significantly only in 'Braeburn' apples. Apples within the 'Ariane' and 'Braeburn' cultivars showed differences in *juiciness*. Even when differences were not significant, for each cultivar, the apples with the lowest acoustic firmness levels had lower scores for *crunchiness*, *firmness*, *crispness* and *juiciness*, and higher scores for *fondant* except in 'Ariane' apples. For *acidity*, the low and high acoustic firmness levels differed significantly only in the 'Ariane' and 'Pink Lady®' cultivars. Whatever the cultivar, apples with a low acoustic firmness level always showed a lower score for *acidity* and *sweetness* compared to apples with a high acoustic firmness level. Better discrimination within 'Braeburn' apples as compared to discrimination within 'Pink Lady®' apples might be explained by the larger differences between low and high acoustic firmness levels for 'Braeburn' apples ($\Delta = 14.6 \text{ Hz}^2 \text{ g}^{2/3}$) compared to the cultivar 'Pink Lady®' ($\Delta = 8.8 \text{ Hz}^2 \text{ g}^{2/3}$). For 'Ariane' apples, the difference between low and high acoustic firmness levels was similar to that for 'Braeburn' apples ($\Delta = 15.6 \text{ Hz}^2 \text{ g}^{2/3}$). In this case, the lower discrimination for 'Ariane' apples might be due to a dependence of the acoustic firmness measurement on cultivar as observed by Molina-Delgado et al. (2009b) for Magness–Taylor firmness and acoustic firmness measurements. As a consequence, for different cultivars, equal ranges of acoustic firmness might not imply equal differences in the sensory perception.

3.3. Discrimination between cultivars, each with low, medium and high acoustic firmness: exp III

Discrimination between apples with different acoustic firmness levels found in experiment II were confirmed. With respect to the previous results, the cultivars differed in all attributes (significant cultivar effect). The ANOVA results showed that assessors were able to discriminate between fruit with different acoustic firmness levels for *crunchiness*, *firmness*, *juiciness* and *acidity* (Table 7). Discrimination between the acoustic firmness levels depended on the cultivar for *sweetness* because the interaction between cultivar and acoustic firmness was significant. Moreover, the post hoc

Table 5
Significance levels for the sensory evaluation of the three apples cultivars with two acoustic firmness levels.

Attributes	Cultivar	Acoustic firmness	Cultivar × acoustic firmness	Assessor	Assessor × cultivar × acoustic firmness
Crunchiness	<0.001	<0.001	0.035	<0.001	1.000
Firmness	<0.001	0.004	0.037	<0.001	0.115
Crispness	<0.001	0.001	0.015	<0.001	0.102
Juiciness	<0.001	<0.001	0.013	<0.001	1.000
Fondant	<0.001	0.075	0.062	<0.001	0.034
Acidity	<0.001	<0.001	0.072	<0.001	0.006
Sweetness	0.001	0.117	0.981	<0.001	0.085

Table 6
Mean scores of the seven sensory attributes for the three apples cultivars with two acoustic firmness levels.

Attributes	Ariane			Braeburn			Pink Lady®		
	Low acoustic firmness	High acoustic firmness	Cultivar average	Low acoustic firmness	High acoustic firmness	Cultivar average	Low acoustic firmness	High acoustic firmness	Cultivar average
Crunchiness	5.43a ^a	5.81a	5.62B ^b	3.93a	5.12b	4.53A	7.82a	8.10a	7.96C
Firmness	4.64a	4.68a	4.66B	2.84a	3.83b	3.34A	7.29a	7.58a	7.44C
Crispness	4.86a	5.26a	5.06B	4.11a	5.19b	4.65A	7.31a	7.36a	7.34C
Juiciness	4.42a	5.43b	4.93A	4.32a	5.78b	5.05A	6.94a	7.22a	7.08B
Fondant	4.87a	4.94a	4.91B	6.80b	5.77a	6.29C	1.79a	1.64a	1.72A
Acidity	4.06a	4.92b	4.49C	2.77a	3.13a	2.95A	3.17a	4.55b	3.96B
Sweetness	5.14a	5.38a	5.26B	4.32a	4.56a	4.44A	4.85a	5.16a	5.01A

^a Values followed by the same small letters within a cultivar are not significantly different ($\alpha = 0.05$).

^b Values followed by a capital letter within a row are not significantly different ($\alpha = 0.05$).

Table 7
Significance levels for the sensory evaluation of the three apple cultivars with three acoustic firmness levels.

Attributes	Cultivar	Acoustic firmness	Cultivar × acoustic firmness	Assessor	Assessor × cultivar × acoustic firmness
Crunchiness	<0.001	0.002	0.661	<0.001	1.000
Firmness	<0.001	0.028	0.617	<0.001	1.000
Crispness	<0.001	0.407	0.535	<0.001	0.349
Juiciness	<0.001	<0.001	0.130	<0.001	0.043
Fondant	<0.001	0.353	0.736	<0.001	0.257
Acidity	<0.001	<0.001	0.305	<0.001	0.069
Sweetness	0.011	0.826	0.010	<0.001	0.998

analysis showed that differences between the acoustic firmness levels were observed within the 'Braeburn' cultivar for *crunchiness*, *juiciness*, *acidity* and *sweetness*; differences within the 'Ariane' cultivar were observed for *juiciness* and *acidity*; and within the 'Pink Lady®' cultivar, the only differences were observed for *acidity*. Furthermore, trends of discrimination were observed amongst the 'Ariane' apples for *crunchiness*; amongst the 'Braeburn' apples for *firmness* and *fondant*, and amongst the 'Pink Lady®' apples for *crunchiness*, *firmness*, *juiciness* and *sweetness*. Only two replicates were performed because of time constraints, which could explain a

smaller discrimination compared to experiment II. For *crunchiness*, *firmness*, *crispness* and *juiciness*, even though the acoustic firmness levels were not well distinguished, the least-firm fruit had lower scores (except for *crispness* in the 'Pink Lady®' apples). For *fondant* the firmest fruit had lower scores. Again, the *acidity* scores were lower for apples with lower acoustic firmness levels, irrespective of the cultivar. More mature apples present lower acoustic firmness (Mehinagic et al., 2006) and lower acid content (Ackermann et al., 1992; Plotto et al., 1999). Therefore, apples with lower acoustic firmness might be less acidic (Table 8).

Table 8
Mean scores of the seven sensory attributes for the three apple cultivars with three acoustic firmness levels.

Attributes	Ariane				Braeburn				Pink Lady®			
	Low acoustic firmness	Medium acoustic firmness	High acoustic firmness	Cultivar average	Low acoustic firmness	Medium acoustic firmness	High acoustic firmness	Cultivar average	Low acoustic firmness	Medium acoustic firmness	High acoustic firmness	Cultivar average
Crunchiness	5.33a ^a (a) ^b	5.96a(b)	6.05a(b)	5.78B ^c	3.99a	4.87b	4.81b	4.56A	7.59a(a)	7.72a(ab)	8.27a(b)	7.86C
Firmness	4.68a	4.80a	5.28a	4.92B	2.73a(a)	3.39a(b)	3.24a(ab)	3.12A	7.13a(a)	7.18a(a)	7.81a(b)	7.37C
Crispness	5.17a	5.40a	5.59a	5.39B	4.32a	4.86a	4.41a	4.53A	7.08a	6.97a	7.42a	7.16C
Juiciness	4.36a	4.75a	5.68b	4.93A	4.09a	5.81b	5.72b	5.21A	6.88a(a)	7.29a(ab)	7.65a(b)	7.27B
Fondant	4.39a	4.60a	4.20a	4.40B	7.27a(b)	6.69a(ab)	6.46a(a)	6.81C	1.94a	1.88a	1.78a	1.87A
Acidity	3.39a	4.33b	5.44c	4.39C	2.74a	3.57b	3.54ab	3.28A	2.61a	3.64b	4.26b	3.50B
Sweetness	5.04a	5.23a	4.99a	5.09A	4.52a	4.26a	5.35b	4.71A	5.38a(ab)	5.68a(b)	4.98a(a)	5.35A

^a Values followed by the same small letters within a cultivar are not significantly different at a 5% level.

^b Values followed by the same small letters in brackets within a cultivar are not significantly different at a 10% level.

^c Values followed by a capital letter within a row are not significantly different at a 5% level.

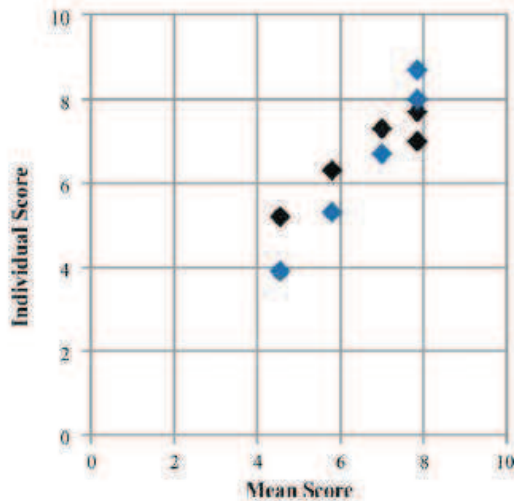


Fig. 1. Individual *crunchiness* scores against mean scores for the five 'Ariane' apples tasted by assessors 471 \blacklozenge and 482 \blacklozenge .

3.4. Including batch variability in the model

The present performance results confirmed the hypothesis that a panel is able to sense small differences between apples of a single cultivar. Therefore, accounting for within-batch variability in the statistical analysis becomes important. To take within-batch variability into account, data from experiment III were analysed without accounting for acoustic firmness levels with the standard ANOVA model (Eq. (5), Section 2) and the ANOVA model including the fruit factor (Eq. (4), Section 2).

First, the scores of the assessors who had tasted the same apple were observed. Differences between the mean scores of the apples of a single cultivar can be larger than differences between the scores given by several assessors for a single apple. For example, assessors 471 and 484 shared the same apples for five of the six 'Ariane' replicates. Regarding the attribute *crunchiness* (Fig. 1), the maximum difference between the scores of the two assessors was 1.7, and the mean difference was 0.9, whereas the maximum differences between the mean scores of the five apples was 3.3. It was noticeable that one fruit was particularly less crunchy compared to the others. This example illustrates the within-batch variability that can occur. The model used until now to measure differences between samples or to evaluate panel performances did not allow

Table 9

Significance levels from the ANOVA model including a random fruit effect for the sensory evaluation of the three apple varieties.

Attributes	Cultivar	Assessor	Assessor \times cultivar	Fruit
Crunchiness	<0.001	<0.001	0.001	<0.001
Firmness	<0.001	<0.001	0.001	<0.001
Crispness	<0.001	<0.001	<0.001	<0.001
Juiciness	<0.001	0.004	<0.001	<0.001
Fondant	<0.001	<0.001	<0.001	<0.001
Acidity	0.001	<0.001	0.974	<0.001
Sweetness	0.012	<0.001	0.178	0.505

for the quantification of fruit contribution in the variability of the scores.

A mixed hierarchical ANOVA that included a fruit effect was applied and showed a significant fruit effect for all sensory attributes except sweetness (Table 9). The fruit, thus, appeared to be an important factor, as found by Aguirre et al. (2008) for mushroom quality obtained with instrumental measurements. The part of the fruit in the variance component was non-negligible. In fact, when a significant effect of fruit was observed, the fruit term in the variance component was always larger than the cultivar-by-assessor term. This result means that the error present in the data was more likely due to variability in the apples than to variability among the assessors. Fig. 2 shows one of the attributes that exhibits a significant fruit effect: *crunchiness*. The contribution of fruit to variance is larger than any other contribution.

To evaluate the impact of the addition of the fruit effect to the model, the two models were compared using F statistics for the fixed cultivar effect and the variance structure for random effects (Table 10). A significant fruit effect led to a decrease in the residual part of the variance and in the assessor-by-cultivar interaction term in the variance structure (except for *fondant*). In addition, the F statistics of the cultivar effect decreased, except for *juiciness*, *acidity* and *sweetness*. The example of *crunchiness* (Fig. 2) shows that the addition of the fruit effect implied a decrease in both the residual and the cultivar-by-assessor interaction terms of the variance. If the fruit effect was to be very large, then the cultivar effect could become insignificant and differences between the cultivars could not be affirmed. This phenomenon was observed by Casals et al. (2011) for some agricultural traits of tomatoes, where, by hierarchical ANOVA, a great variability within the cultivars was charged for no differences between cultivars. Moreover, for the sensory quality, they applied a model without taking the fruit effect into account and found a product effect due to intra-cultivar variability and not due to differences between cultivars. In this case, the

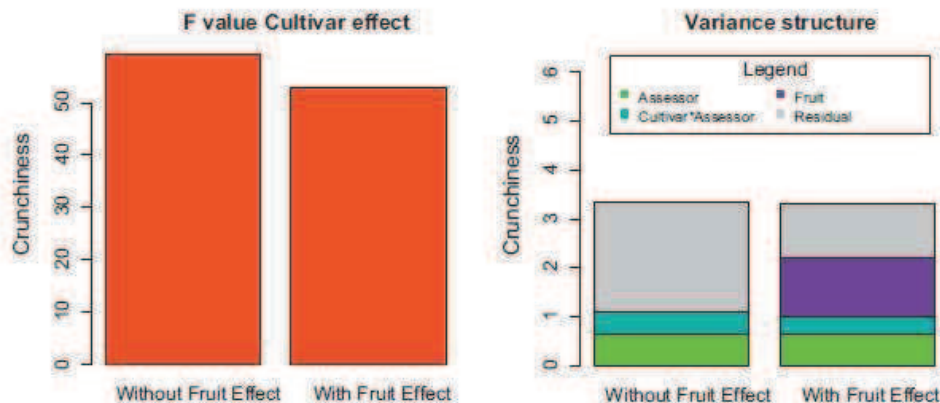


Fig. 2. Evolution of ANOVA results with fruit effect inclusion for *crunchiness*: F statistics for the cultivar effect (left) and variance structure for the random effects (right).

Table 10
F value for the fixed effect and variance component for the random effect of the two models.

Fruit included	Cultivar		Assessor		Assessor × cultivar		Fruit		Residual	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Crunchiness	59.23	52.89	0.62	0.65	0.47	0.34	–	1.22	2.26	1.10
Firmness	87.49	83.11	0.69	0.79	0.55	0.38	–	1.17	2.32	1.26
Crispness	36.66	33.09	0.58	0.78	0.54	0.44	–	0.94	2.04	1.16
Juiciness	32.21	24.42	0.46	0.57	0.45	0.51	–	1.49	2.80	1.32
Fondant	73.80	75.48	1.17	1.22	1.05	0.79	–	1.07	2.66	1.70
Acidity	10.64	7.25	2.42	2.39	0.00	0.04	–	1.39	3.47	2.10
Sweetness	3.22	2.93	1.63	1.63	0.18	0.21	–	0.15	2.34	2.18

application of hierarchical ANOVA that included a fruit effect for sensory characteristics could have led to conclusions similar to those on agricultural traits and emphasised the large intra-cultivar variability.

4. Conclusion

This work confirmed the need for the adaptation of the methodology for sensory quality evaluation of postharvest products. The trained panel was effective in determining the differences between apples within a cultivar, even in an experimental plan that presented larger differences due to cultivar differences. In consequence, data analysis for sensory quality evaluation of apple has been adapted and a mixed hierarchical ANOVA that included a fruit effect proved to be effective in accounting for within-batch variability. Within-batch variability was quantified and identified as an important source of the total sensory scores variability. This variability was even larger than the assessor's differences.

Thus, for sensory quality evaluation of apples, we recommend (1) to share each tested apple among several assessors and (2) to use of mixed hierarchical ANOVA when exploring and interpreting sensory scores in order to be able to quantify within-batch variability. It should be noticed that sharing apples is a requirement which allows for correct data analysis.

Finally, using the mixed hierarchical ANOVA to analyse apples sensory scores leads to an increase in understanding of the important sources of variability in the data. Therefore, the proposed methodology helps to better comprehend the postharvest quality of apples.

Acknowledgements

The authors thank Dr Mara V. Galmarini for her constructive remarks and advice, Corinne Patron and Isabel Saillard for their assistance and the assessors who participated in the descriptive sensory analysis. This research was funded by the Conseil Régional des Pays de Loire.

References

- Abbott, J., Lu, R., 1996. Anisotropic mechanical properties of apples. *Transaction of the ASAE* 39, 1451–1459.
- Abbott, J.A., Affeldt, H.A., Liljedahl, L.A., 1992. Firmness measurement of stored 'delicious' apples by sensory methods, Magness–Taylor, and sonic transmission. *Journal of the American Society for Horticultural Science* 117, 590–595.
- Abbott, J.A., Saftner, R.A., Gross, K.C., Vinyard, B.T., Janick, J., 2004. Consumer evaluation and quality measurement of fresh-cut slices of 'Fuji', 'Golden Delicious', 'Gold Rush', and 'Granny Smith' apples. *Postharvest Biology and Technology* 33, 127–140.
- Ackermann, J., Fischer, M., Amado, R., 1992. Changes in sugars, acids, and amino acids during ripening and storage of apples (cv. Glockenapfel). *Journal of Agricultural and Food Chemistry* 40, 1131–1134.
- Aguirre, L., Frias, J.M., Barry-Ryan, C., Grogan, H., 2008. Assessing the effect of product variability on the management of the quality of mushrooms (*Agaricus bisporus*). *Postharvest Biology and Technology* 49, 247–254.
- Barreiro, P., Ortiz, C., Altisent, M.R., Schotte, S., Andani, Z., Wakeling, I., Beyt, P.K., 1998. Comparison between sensory and instrumental measurements of firmness assessment in apples. A collaborative test. *Journal of Texture Studies* 29, 509–525.
- Bhagwat, A.A., Saftner, R.A., Abbott, J.A., 2004. Evaluation of wash treatments for survival of foodborne pathogens and maintenance of quality characteristics of fresh-cut apple slices. *Food Microbiology* 21, 319–326.
- Bobelyn, E., Serban, A.S., Nicu, M., Lammertyn, J., Nicolai, B.M., Saey, W., 2010. Postharvest quality of apple predicted by NIR-spectroscopy: study of the effect of biological variability on spectra and model performance. *Postharvest Biology and Technology* 55, 133–143.
- Brockhoff, P.B., 2003. Statistical testing of individual differences in sensory profiling. *Food Quality and Preference* 14, 425–434.
- Brockhoff, P.B., Skovgaard, I.M., 1994. Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference* 5, 215–224.
- Brookfield, P.L., Nicoll, S., Gunson, F.A., Harker, F.R., Wohlers, M., 2011. Sensory evaluation by small postharvest teams and the relationship with instrumental measurements of apple texture. *Postharvest Biology and Technology* 59, 179–186.
- Casals, J., Pascual, L., Cañizares, J., Cebolla-Cornejo, J., Casañas, F., Nuez, F., 2011. The risks of success in quality vegetable markets: possible genetic erosion in Marmande tomatoes (*Solanum lycopersicum* L.) and consumer dissatisfaction. *Scientia Horticulturae* 130, 78–84.
- Chen, H., DeBaerdemaeker, J., 1993. Finite-element-based modal analysis of fruit firmness. *Transaction of the ASAE* 36, 1827–1834.
- Chen, H., DeBaerdemaeker, J., 1993. Effect of apple shape on acoustic measurements of firmness. *Journal of Agricultural Engineering Research* 56, 253–266.
- Dailant-Spinnier, B., MacFie, H.J.H., Beyts, P.K., Hedderley, D., 1996. Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference* 7, 113–126.
- De Belie, N., Schotte, S., Coucke, P., De Baerdemaeker, J., 2000. Development of an automated monitoring device to quantify changes in firmness of apples during storage. *Postharvest Biology and Technology* 18, 1–8.
- De Ketelaere, B., Stulens, J., Lammertyn, J., Cuong, N.V., De Baerdemaeker, J., 2006. A methodological approach for the identification and quantification of sources of biological variance in postharvest research. *Postharvest Biology and Technology* 39, 1–9.
- Findlay, C.J., Castura, J.C., Lesschaeve, I., 2007. Feedback calibration: a training method for descriptive panels. *Food Quality and Preference* 18, 321–328.
- Hampson, C.R., Quamme, H.A., Hall, J.W., MacDonald, R.A., King, M.C., Cliff, M.A., 2000. Sensory evaluation as a selection tool in apple breeding. *Euphytica* 111, 79–90.
- Harker, F.R., Anne Gunson, F., Brookfield, P.L., White, A., 2002a. An apple a day: the influence of memory on consumer judgment of quality. *Food Quality and Preference* 13, 173–179.
- Harker, F.R., Gunson, F.A., Jaeger, S.R., 2003. The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples. *Postharvest Biology and Technology* 28, 333–347.
- Harker, F.R., Kupferman, E.M., Marin, A.B., Gunson, F.A., Triggs, C.M., 2008. Eating quality standards for apples based on consumer preferences. *Postharvest Biology and Technology* 50, 70–78.
- Harker, F.R., Maindonald, J., Murray, S.H., Gunson, F.A., Hallett, I.C., Walker, S.B., 2002b. Sensory interpretation of instrumental measurements. 1. texture of apple fruit. *Postharvest Biology and Technology* 24, 225–239.
- Heintz, C.M., Kader, A.A., 1983. Procedures for the sensory evaluation of horticultural crops [Quality, color and texture]. *HortScience* 18, 18–22.
- Hertog, M.L., 2002. The impact of biological variation on postharvest population dynamics. *Postharvest Biology and Technology* 26, 253–263.
- Hertog, M.L., Lammertyn, J., Desmet, M., Scheerlinck, N., 2004. The impact of biological variation on postharvest behaviour of tomato fruit. *Postharvest Biology and Technology* 34, 271–284.
- Ioannides, Y., Howarth, M.S., Raithatha, C., Defernez, M., Kemsley, E.K., Smith, A.C., 2007. Texture analysis of Red Delicious fruit: towards multiple measurements on individual fruit. *Food Quality and Preference* 18, 825–833.
- ISO 8586-2:1994 (1994). Analyse sensorielle – Guide général pour la sélection, l'entraînement et le contrôle des sujets – Partie 2: Experts. Geneva, International Organization for Standardization, pp. 83–97.
- ISO 5495 (2005). Analyse sensorielle – Méthodologie – Essai de comparaison par paires. International Organization for Standardization, pp. 321–343.

- Jaeger, S.R., Andani, Z., Wakeling, I.N., MacFie, H.J.H., 1998. Consumer preferences for fresh and aged apples: a cross-cultural comparison. *Food Quality and Preference* 9, 355–366.
- King, G.J., Maliepaard, C., Lynn, J.R., Alston, F.H., Durel, C.E., Evans, K.M., Griffon, B., Laurens, F., Manganaris, A.G., Schrevens, E., 2000. Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). *Theoretical and Applied Genetics* 100, 1074–1084.
- Kingston, C., 1991. Maturity indices for apple and pear. *Horticultural Reviews* 13, 407–432.
- Konopacka, D., Plochanski, W.J., 2004. Effect of storage conditions on the relationship between apple firmness and texture acceptability. *Postharvest Biology and Technology* 32, 205–211.
- Kuznetsova, A., Christensen, R.H.B., Brockhoff, P.B., 2012. lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 1.0–2.
- Lawless, H.T., Heymann, H., 1998. *Sensory Evaluation of Food: Principles and Practices*. Chapman & Hall, New York.
- Mehinagic, E., Royer, G., Bertrand, D., Symoneaux, R., Laurens, F., Jourjon, F., 2003. Relationship between visible-NIR spectroscopy and sensory analysis of apples. *Food Quality and Preference* 14, 473–484.
- Mehinagic, E., Royer, G., Symoneaux, R., Jourjon, F., 2006. Relationship between apple sensory attributes and instrumental parameters of texture. *Journal of Fruit and Ornamental Plant Research* 14, 25–37.
- Meilgaard, M., Civille, G.V., Carr, B.T., 1991. *Sensory Evaluation Techniques*, 2nd ed. CRC Press, Boca Raton, FL.
- Molina-Delgado, D., Alegre, S., Barreiro, P., Valero, C., Ruiz-Altisent, M., Recasens, I., 2009a. Addressing potential sources of variation in several non-destructive techniques for measuring firmness in apples. *Biosystems Engineering* 104, 33–46.
- Molina-Delgado, D., Alegre, S., Puy, J., Recasens, I., 2009b. Relationship between acoustic firmness and Magness Taylor firmness in Royal Gala and Golden Smoothie apples. *Food Science and Technology International* 15, 31–40.
- Montero-Prado, P., Rodríguez-Lafuente, A., Nerin, C., 2011. Active label-based packaging to extend the shelf-life of "Calanda" peach fruit: changes in fruit quality and enzymatic activity. *Postharvest Biology and Technology* 60, 211–219.
- Næs, T., 1998. Detecting individual differences among assessors and differences among replicates in sensory profiling. *Food Quality and Preference* 9, 107–110.
- Næs, T., Brockhoff, P.B., Tomic, O., 2010. *Statistics for sensory and consumer science*. John Wiley & Sons Ltd, Chichester.
- Oraguzie, N., Alspach, P., Volz, R., Whitworth, C., Ranatunga, C., Weskett, R., Harker, R., 2009. Postharvest assessment of fruit quality parameters in apple using both instruments and an expert panel. *Postharvest Biology and Technology* 52, 279–287.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed-Effects Models in S and S-plus®*. Springer Verlag LLC, New York.
- Plotto, A., McDaniel, M.R., Mattheis, J.P., 1999. Characterization of Gala Apple aroma and flavor: differences between controlled atmosphere and air storage. *Journal of the American Society for Horticultural Science* 124, 416–423.
- R Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>
- Romano, R., Brockhoff, P.B., Hersleth, M., Tomic, O., Næs, T., 2008. Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Quality and Preference* 19, 197–209.
- Schouten, R.E., Jongbloed, G., Tijskens, L.M.M., van Kooten, O., 2004. Batch variability and cultivar keeping quality of cucumber. *Postharvest Biology and Technology* 32, 299–310.
- Seppä, L., Railio, J., Mononen, R., Tahvonen, R., Tuorila, H., 2012. From profiles to practice: communicating the sensory characteristics of apples to the wider audience through simplified descriptive profiles. *LWT – Food Science and Technology* 47, 46–55.
- Stevens, M.A., Albright, M., 1980. An approach to sensory evaluation of horticultural commodities [fresh market tomatoes]. *HortScience* 15, 48–50.
- Symoneaux, R., Galmarini, M.V., Mehinagic, E., 2012. Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference* 24, 59–66.
- Symoneaux, R., Royer, G., Madieta, E., Jourjon, F., 2005. Acoustic and sensory measurements of different ripeness of apples. *Acta Horticulturae* 682, 1475–1480.
- Tijskens, L.M.M., Konopacki, P., 2003. Biological variance in agricultural products. Theoretical considerations. *Acta Horticulturae* 600, 661–670.
- Tijskens, L.M.M., Konopacki, P., Simcic, M., 2003. Biological variance, burden or benefit? *Postharvest Biology and Technology* 27, 15–25.
- Vaysse, P., Landry, P., 2004. *Pomme-poire: de la récolte au conditionnement: outils pratiques*. Editions Centre technique interprofessionnel des fruits et légumes.
- Watada, A.E., Abbott, J.A., Hardenburg, R.E., 1980. Sensory characteristics of apple fruit. *Journal of the American Society for Horticultural Science* 105, 371–375.
- Williams, A.A., Carter, C.S., 1977. A language and procedure for the sensory assessment of Cox's Orange Pippin apples. *Journal of the Science of Food and Agriculture* 28, 1090–1104.

Non destructive sampling to improve sensory descriptive results of varying products : an application on apple

C. Bavay¹, R. Symoneaux¹, J. Maitre¹, P. Courcoux², E. Mehinagic¹

¹LUNAM University, Groupe ESA, GRAPPE [Research Group of Agro-Food Products and Process], France

²LUNAM University, ONIRIS, Sensometrics and Chemometrics Laboratory, France

Contact: c.bavay@groupe-esa.com

Introduction

Despite the training process of a panel, data resulting from sensory profiling of fruit present a large variability: it is due to both individual differences between panellists and fruit heterogeneity related to different physiological and environmental factors. Even different parts of the same fruit may vary. Besides, data processing misses in separating product variability and panel lack of consensus. Many authors tried to predict textural sensory descriptors by instrumental parameters. Among their results, high correlations between textural sensory descriptors of apples, as crunchiness and juiciness, and acoustic firmness index were shown [1].

Objective

The objective of the study was to measure, for three apple cultivars, the impact of reduction of fruit heterogeneity by non-destructive sampling on the reliability of sensory profile results (i.e. dispersion and panel discrimination).

Material and Methods

Samples

Three cultivars presenting close textural sensory characteristics: Ariane (AR), Braeburn (BR) and Pink Lady® (PL) (Fig. 1)

Sampling procedure

(1) Non-destructive measurement of firmness index (Fi), based on the acoustic properties of a large batch of apples.

$$Fi = f_r^2 \cdot m^{2/3}$$

with f_r the first resonance frequency and m the weight of the fruit [2-2];

(2) Sorting by firmness index and division of the batch into five equal range classes;

(3) Three sampling methods to get batches of 20 apples with different degrees of heterogeneity:

→ a heterogeneous batch → a normally distributed batch → a homogeneous batch

Sampling methods were centred on the initial batch mean resulting in equal means for the three new batches. (Fig. 2)

Descriptive sensory analysis

→ Nineteen trained panellists experienced in apple tasting;

→ Seven descriptors ('Crunchy' (C), 'Firm' (F), 'Crispy' (Cr), 'Juicy' (J), 'Fondant' (Fo), 'Acid' (A), 'Sweet' (S)) evaluated on an unstructured scale (0=not, 10=extremely);

→ Three replicates.

Statistical analysis

R software was used for Levene tests (comparison of more than two variances) and analysis of variance.



Fig. 1. Sensory profiles of the three chosen cultivars.

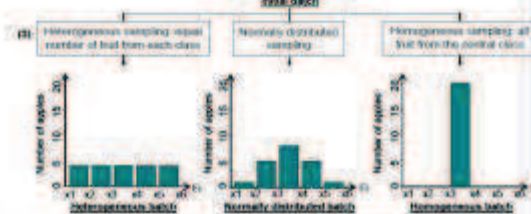
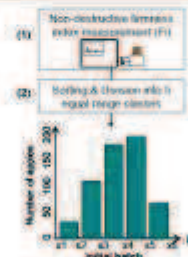


Fig. 2. Three sampling procedures applied in an initial batch to obtain batches with different degrees of heterogeneity.

Results

Intra-cultivar variance

Results vary according to cultivars:

→ Ariane & Braeburn: Intra-cultivar sensory variances do not decrease from heterogeneous to homogeneous batch (Fig. 3);

→ Pink lady®: Intra-cultivar sensory variances decrease from heterogeneous to normally distributed batch for 'Crunchy', 'Firm', 'Crispy' and 'Fondant' (Fig. 3). Variance values for normally distributed and homogeneous batches are close.

Levene tests (Table 1) show:

→ Ariane: no significant differences between batches from different sampling methods;

→ Braeburn: decreases of variance are significant for normally distributed batch for 'Juicy' and 'Fondant';

→ Pink Lady®: decreases of variance are significant for normally distributed and homogeneous batches for 'Crunchy' and 'Crispy'.

Discrimination

Whatever the sampling method, the panel discriminate cultivars (p-value < 0.001). Discrimination is not improved by homogenisation with firmness index except for 'Acid' (Table 2). The link between sensory parameters and firmness index could be challenged; according to these results.

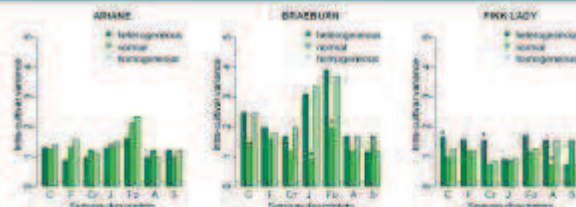


Fig. 3. Intra-cultivar variances of Ariane, Braeburn and Pink Lady® (from left to right) for the three samplings and each descriptor.

Table 1. Levene test results – F value.

	AR	BR	PL
C	0.086	0.969	2.191*
F	1.228	0.641	0.179
Cr	0.368	0.821	2.718*
J	0.202	11.591*	0.130
Fo	1.485	1.139*	3.002
A	0.661	0.276	2.561*
S	0.967	0.738	0.256

Table 2. Discrimination levels – F value for Cultivar effect from ANOVA.

	C	F	Cr	J	Fo	A	S
Heterogeneous	22.58	37.23	11.68	4.70	33.47	3.57	19.72
Normally distributed	20.47	30.71	8.30	11.22	21.68	14.40	20.49
Homogeneous	16.49	25.68	4.41	6.33	18.66	22.25	16.31

*Signif. index: 90.1

Conclusion & Perspectives

The homogenisation by non destructive acoustic sampling implies a significant decrease of variance for sensory descriptors for Pink Lady®. These promising results encourage the investigation of sampling methods to improve sensory results reliability. Nonetheless, the acoustic firmness index does not seem to be the best measure to efficiently reduce the sensory data variance. The reliability of this tool in terms of link with sensory data has already been raised [3]. Other non destructive measures, as optical properties, should be considered. Beyond that, a recent experimentation using Magness-taylor index as sampling measure seems more promising.

References

- Barreiro P, Ordo C, Allison M B, Schote S, Andari Z, Waking I, Beyl P S. (1998). Comparison between sensory and instrumental measurements for maturity assessment in apples. A collaborative test. Journal of Texture Studies, 29(5), 509-525.
- Coste, J. B. and Baud R.H. (1973). A mathematical study of resonance in intact fruits and vegetables using a 3-media elastic sphere model. Journal of Agricultural Engineering Research 18: 1417-157.
- Mehinagic E., Rayer G., Symoneaux R., Inarjen F. (2006). Objective measurements of apple texture with geometrometry, compression and acoustic methods in relation to the sensory perceptions. Proceedings of the 8th International Symposium on Food Texture and Structure, Zurich, Switzerland.

Acknowledgement

The authors thank Corinne Patron for her aid, and the panellists who did the descriptive sensory analysis. C. Bavay thanks the Conseil Régional des Pays de la Loire for their funding.



9th Pangborn Sensory Science Symposium
4-8 September 2011
The Sheraton Centre Toronto Hotel, Toronto, Canada





Management of fruit presentation in sensory evaluation of apples for more reliable results

C. Bavay¹, R. Symoneaux¹, I. Maitre¹, E. Mehinagic¹

¹ LUNAM University, Groupe ESA, UPSP GRAPPE, France
Contact: c.bavay@groupe-esa.com



Introduction

Sensory analysis is used for detection of very sensitive differences between close-varieties or even clones. Despite the training process of a panel, data resulting from sensory profiling of fruit present a large variability. It is due to both individual differences between assessors and fruit heterogeneity. Even different parts of the same fruit may vary. One of the ways to improve the consensus between assessors and their global performance is to reduce sample heterogeneity using an adequate sampling method. [1] [2]

Objective

The objective of the present study was to develop a suitable sampling method for the sensory evaluation of heterogeneous raw material to get more reliable sensory results.

Material and Methods

Samples

Three cultivars : Ariane, Braeburn and Pink Lady®.

Sampling procedure

Three sampling methods were applied to simulate a decreasing intra-variability of the tasted sample: an apple per assessor (M1), an apple for four assessors (M2) and an apple for eight assessors (M3).



Descriptive sensory analysis

- Fifteen trained panelists trained in apple tasting during 130 h;
- Five descriptors : Crunchiness (C), Firmness (F), Juiciness (J), Fondant (F) and Acidity (A); evaluated on an unstructured scale (0=not, 10=extremely);
- Four replicates.

Statistical analysis

- Variances were compared with Cochran test
- Hierarchical mixed analysis of variance was applied to evaluate the performances of the panel
- R software was used

$$X_{ijk} = \text{Cultivar}_i + \text{Assessor}_j + \text{Cultivar: Assessor}_{ij} + \epsilon_{ijk}$$

With $\epsilon_{ijk} \sim N(0, \sigma^2)$, $\text{Assessor}_j \sim N(0, \sigma^2_{\text{assessor}})$ et $\text{Cultivar: Assessor}_{ij} \sim N(0, \sigma^2_{\text{cultivar: assessor}})$, and all terms are independent.

Results

Intra-cultivar variance

Results vary according to cultivars. The intra-cultivar variances are generally similar for the three methods, A decrease can be observed when each apple is shared by several assessors (M2 and M3).

Discrimination

M2: Compared with M1 results, discrimination between cultivars is improved for *Crunchiness, Firmness, Fondant and Acidity* (Fig. 1).

M3: Compared with M1 results, discrimination between cultivars is improved for *Firmness, Juiciness, Fondant and Acidity* (Fig. 1).

In addition, reproducibility is better with method M2 than method M3 (Fig. 2.).

F value discrimination

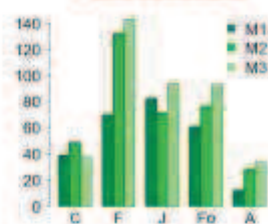


Fig. 1. F value for Cultivar effect

Residual mean square



Fig. 2. Residual mean square

Conclusion & Perspectives

This results show that sample presentation has an impact on the reliability of sensory results. In this case, method M2 seems to best respond to our objective. Through a little change in methodology, discrimination is improved. This improvement is due to a decrease of fruit-to-fruit variability implying an increase of agreement at list among assessors tasting the same apple. Therefore, our recommendations will be to have two to four assessors tasting the same apple.

References

- [1] Heintz, C.M., & Kader, A.A., (1983). Procedures for the sensory evaluation of horticultural crops [Quality, color and texture]. HortScience (USA), 18 (1), 18-22.
- [2] Abbott, J.A., Saftner, R.A., Gross, K.C., Vinyard, B.T., & Janick, J., (2004). Consumer evaluation and quality measurement of fresh-cut slices of 'Fuji', 'Golden Delicious', 'GoldRush', and 'Granny Smith' apples. Postharvest biology and technology, 33 (2), 127-140.

Acknowledgement

The authors thank Corinne Patron and Isabel Saillard for their assistance, and the panelists who did the descriptive sensory analysis. This research was funded by the Conseil Régional des Pays de Loire.



SYMPOSIUM ON HORTICULTURE IN EUROPE





Sensory profile adaptation to biological variability of fruit batch

C. Bavay¹, R. Symoneaux¹, I. Maitre¹, E. Mehinagic¹

¹ LUNAM University, Groupe ESA, UPSP GRAPPE, France
Contact: c.bavay@groupe-esa.com



Introduction

Sensory analysis is used for detection of very sensitive differences between close varieties or even clones. Despite the training process of a panel, data resulting from sensory profiling of fruit present a large variability. It is due to both individual differences between assessors and fruit heterogeneity. Even different parts of the same fruit may vary. One of the ways to improve the consensus between assessors and their global performance is to adapt the tasting procedure to enable the quantification of sample variability.

Objective

The objective of the present study was to compare results from two models of analysis of variance (ANOVA) and look into consistency gain when including a factor that takes biological variability into account.

Material & Methods

Descriptive Sensory analysis

- 3 cultivars
- 19 trained assessors divided in 6 sub-panels (3-4)
- 7 descriptors : *Crunchiness, Firmness, Crispness, Juiciness, Fondant, Acidity, Sweetness*; evaluated on an unstructured scale (0=not, 10=extremely)
- 3 replicates

Tasting procedure

Each apple is divided into 4 portions, in that way each individual apple is scored by several assessors



Statistical analysis

Main objectives of ANOVA: identify and quantify the factors responsible for the variability of the response

Standard model

$$Y_{ijk} = \text{Culti var}_i + \text{Assessor}_j + \text{Culti var}:\text{Assessor}_{ij} + \varepsilon_{ijk} \quad (1)$$

Model that takes biological variability into account

$$Y_{ijk} = \text{Culti var}_i + \text{Assessor}_j + \text{Cultivar}:\text{Assessor}_{ij} + \text{Fruit}_{ij} + \varepsilon_{ijk} \quad (2)$$

Results

The fruit contribution is major

- Larger than the assessor contributions
- Larger than the cultivar contribution (e.g. for juiciness)



Fig. 1. Contribution to variance for crunchiness and juiciness [%]

Conclusions about product differences may change

- The discrimination decreases when the biological variability is taken into account

	Model (1)	Model (2)
Crunchiness	1.02E-14	2.89E-07
Firmness	2.76E-19	1.11E-09
Crispness	3.00E-06	3.00E-06
Juiciness	1.02E-03	6.56E-02
Fondant	4.85E-15	3.18E-06
Acidity	3.20E-03	4.72E-02
Sweetness	2.58E-09	2.58E-09

Signif. code:
■ p < 0.05
■ p < 0.01
■ p < 0.001

Conclusion & Perspectives

The fruit contribution to variance is quantified. It appears to be an important factor of variability. Including a factor to take the biological variability existing within each sample in the ANOVA is relevant, and is possible if and only if each sample is tasted by several assessors. So, training, performance and profile measurement should be run with several assessors tasting the same sample. The panel leader will therefore correctly evaluate assessors' aptitude, measure the sample contribution to the variability of the response, and draw more reliable conclusions. As a perspective the model proposed to take biological variability into account may be combined with Mixed Assessor Model [1]. This procedure might be extended to any product presenting biological variability, such as vegetables, dairy products, cheeses...

References

[1] Brockhoff, P.B., & Skovgaard, I.M., (1994). Modelling individual differences between assessors in sensory evaluations. Food Quality and Preference, 5 (3), 215-224.

Acknowledgement

The authors thank Corinné Patron and Isabel Saillard for their assistance, and the panellists who did the descriptive sensory analysis. This research was funded by the Conseil Régional des Pays de Loire.



The importance of methodological adaptation to get reliable results in sensory evaluation of product presenting biological variability:



a case study on apple

C. Bavay¹, R. Symoneaux¹, I. Maitre¹, A. Kuznetsova², P.B. Brockhoff¹, E. Mehinagic¹

¹ LUNAM University, Groupe ESA, UPSP GRAPPE, France

² DTU Informatics, Statistical section, Technical University of Denmark, Denmark

Contact: c.bavay@groupe-esa.com

Introduction

Sensory profiling is a strategic tool to manage product quality and precise what is expected by consumers. A standard analysis of variance (ANOVA), including fixed product effect, random assessor effect and interaction (random) is often used to analyze sensory data. However this model may not be the best to handle all types of data. For example, in sensory evaluation of product presenting biological variability, the variability of the response is due to both individual differences between assessors and product heterogeneity.

Objective

The objective of the present study was to compare results from two models of analysis of variance and look into consistency gain when including a factor that takes biological variability into account.

Material & Methods

Descriptive Sensory analysis

- 3 cultivars (18 apples per cultivar)
- 19 trained assessors divided in 6 sub-panels (3-4)
- 7 descriptors : *Crunchiness, Firmness, Crispness, Juiciness, Fondant, Acidity & Sweetness*; evaluated on an unstructured scale (0=not, 10=extremely)
- 3 replicates

Tasting procedure

Each apple is divided into 4 portions, in that way each individual apple is scored by several assessors



Statistical analysis

Main objectives of ANOVA: identify and quantify the factors responsible for the variability of the response

Standard model

$$Y_{ijk} = \text{Culti var}_i + \text{Assessor}_j + \text{Culti var}:\text{Assessor}_{ij} + \epsilon_{ijk} \quad (1)$$

Model that takes biological variability into account

$$Y_{ijk} = \text{Culti var}_i + \text{Assessor}_j + \text{Culti var}:\text{Assessor}_{ij} + \text{Fruit}_{l(i)} + \epsilon_{ijk} \quad (2)$$

Analyses were run with `lme4`Test package in R 2.15.1 [1]

Results

The fruit contribution is major

- Larger than the assessor contributions
- Larger than the cultivar contribution (e.g. for juiciness)

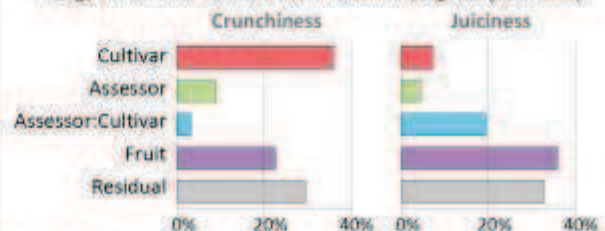


Fig. 1. Contribution to variance for crunchiness and juiciness [%]

Conclusions about product differences may change

- The discrimination decreases when the biological variability is taken into account

Table 1. p value of the two model situations for cultivar effect

Signif. code:
■ p < 0.05
■ p < 0.01
■ p < 0.001

	Model (1)	Model (2)
Crunchiness	1.02E-14	2.89E-07
Firmness	2.76E-19	1.11E-09
Crispness	3.00E-06	3.00E-06
Juiciness	1.02E-03	6.56E-02
Fondant	4.85E-13	3.18E-08
Acidity	3.20E-03	4.72E-02
Sweetness	2.58E-09	2.58E-09

Conclusion & Perspectives

The fruit contribution to variance appears to be an important factor of variability. Including a factor to take the biological variability existing within each sample in the ANOVA is relevant, and is possible if and only if each sample is tasted by several assessors. So, for the characterization of products presenting biological variability, samples should be shared by several assessors in order to provide more reliable sensory results. The adaptation of sensory protocols and data analysis are here pointed out as key-points for the reliability of sensory results. As a perspective the model proposed to take biological variability into account may be combined with Mixed Assessor Model [2]. This procedure might be extended to any product presenting biological variability, such as vegetables, dairy products, cheeses...

- References** [1] Kuznetsova, A., Christensen, R.H.B., Bavay, C., Brockhoff, P.B., (2012). Automated Mixed ANOVA Modelling of sensory and consumer data. In 11th Sensometrics, 10-13 July 2012, Rennes, France
 [2] Brockhoff, P.B., & Skovgaard, I.M., (1994). Modelling individual differences between assessors in sensory evaluations. Food Quality and Preference, 5 (3), 215-224.

Acknowledgement

The authors thank Corinne Patron and Isabel Saillard for their assistance, and the panellists who did the descriptive sensory analysis. This research was funded by the Conseil Régional des Pays de Loire.



Thèse de Doctorat

Cécile Bavay

Adaptation des méthodologies d'évaluation sensorielle aux produits agroalimentaires à forte variabilité

Résumé

Les caractéristiques sensorielles peuvent être mesurées par des méthodes descriptives d'évaluation sensorielle, parmi lesquelles le profil conventionnel. Les résultats sensoriels de profil présentent une forte variabilité, due, d'une part aux différences entre les sujets du panel, et d'autre part à la variabilité intra-lot, et notamment aux différences existant entre les individus d'un échantillon. À ce jour, la variabilité des sujets est prise en compte et les méthodologies de collecte et d'analyse des données ont été adaptées dans ce but. Néanmoins, la variabilité intra-lot n'a pas été abordée, malgré le challenge qu'elle représente. Avec pour produit modèle la pomme, ce travail de thèse aborde la problématique de la variabilité intra-lot sous deux angles : la mesure de la variabilité et la réduction de la variabilité, dans le but d'obtenir une mesure sensorielle plus fiable.

Premièrement, la capacité de discrimination du panel a été observée entre les variétés étudiées et entre les fruits de chaque variété, identifiés comme différents selon une mesure instrumentale. Deuxièmement, afin de prendre en compte la variabilité intra-lot, le modèle d'analyse de la variance classiquement utilisé en évaluation sensorielle a été adapté avec l'ajout d'un facteur *fruit* hiérarchisé au facteur *variété*. Cette adaptation implique le partage de chaque fruit par plusieurs sujets. L'application de ce modèle a démontré l'importance de la variabilité intra-lot et l'importance du choix du modèle pour obtenir des résultats plus pertinents. De plus, le modèle d'analyse de la variance déterminé a été adapté afin de distinguer désaccord pur entre les sujets et différences de dispersion. Enfin, une réduction de la variabilité effective est obtenue par le partage des fruits ou par une homogénéisation instrumentale, cette dernière étant dépendante de la variété et de la mesure choisie.

Les résultats de ce travail apportent des clés méthodologiques pour l'obtention de résultats fiables dans la cadre de l'évaluation sensorielle de produits variables.

Mots clés

analyse sensorielle, analyse de la variance, fruits et légumes, modèle mixte hiérarchique, profil conventionnel, performances, pomme, sensométrie, variabilité

Abstract

The sensory characteristics are evaluated by descriptive methods of sensory evaluation, such as the conventional profile. The sensory results from profiling present a large variability, due to differences between assessors on the one hand, and within-batch variability, that is differences between units of a sample, on the other hand. To date, differences between assessors are taken into account and sensory methodologies have been adapted for that. Nevertheless, within-batch variability has not been tackled, despite the challenge it represents. With the apple as a model, this PhD work addresses the issue of within-batch variability from two points of view: the measuring of variability and the reduction of variability, with the aim of improving the reliability of the sensory measure.

First, the capacity of the panel to discriminate between apples from different cultivars and between apples, identified as different by an instrumental measure, within each cultivar has been observed. Second, in order to take within-batch variability into account, the standard models for analysis of variance used in sensory evaluation has been adapted by adding a *fruit* factor nested within the *cultivar* factor. This adaptation requires the sharing of each piece of fruit by several assessors. The application of this model has demonstrated the importance of within-batch variability and the consequence of model choice to obtain more relevant results. In addition, the determined model for analysis of variance has been adapted to distinguish between assessors' disagreement and scaling differences. Finally, a reduction of variability has been obtained through the sharing of fruits and after the homogenization by instrumental measurement, the latter depending on the cultivar and on the chosen measure.

The results of the present work provide methodological keys to obtain reliable sensory results in the frame of the evaluation of products subject to biological variation.

Key Words

analysis of variance, apple, conventional profile, sensory analysis, fruit and vegetables, hierarchical mixed modelling, performances, sensometrics, variability