

Emilien Ghomi – Equipe in|situ| - LRI

*“Designing expressive interaction techniques for novices inspired by expert activities:
the case of musical practice”*

*“La conception de techniques d’interaction expressives adaptées aux novices et inspirées
par des activités expertes : le cas de la pratique musicale”*

Introduction

Les systèmes informatiques étant de plus en plus complexes, notre rôle en tant que concepteurs de systèmes interactifs est de proposer des méthodes efficaces et pérennes pour créer des techniques d’interaction qui soient à la fois expressives, efficaces et utilisables.

Les designers considèrent souvent que, pour une activité donnée, seuls les experts sont capables d’atteindre de hauts niveaux d’expressivité, après des années de pratique. Notre approche, basée sur la littérature en phénoménologie et en psychologie, montre que les non-experts ont une connaissance considérable d’activités expertes ayant une portée sociale, comme les activités artistiques, et qu’ils l’acquièrent de façon implicite.

En effet, les non-musiciens connaissent implicitement les règles liées au rythme, à l’harmonie et à la tonalité de la musique tonale occidentale. Les concepteurs d’interaction peuvent réutiliser cette connaissance pour créer des techniques d’interaction expressives utilisables par des novices. Nous proposons également d’utiliser les artefacts experts comme source d’inspiration car ils sont efficaces dans des tâches complexes, et de s’inspirer des méthodes d’apprentissage expertes pour aider les utilisateurs à progresser rapidement dans leur apprentissage de ces techniques d’interaction.

Aujourd’hui, un important courant de la conception de systèmes interactifs défend l’idée que la relation entre l’humain et l’ordinateur doit être « simple », « intuitive » voire « naturelle ». Cette approche a le mérite d’insister sur le fait que l’engagement de l’utilisateur doit être peu coûteux à la fois physiquement et intellectuellement, mais semble s’apparenter à l’illusion répandue lors de l’arrivée des instruments de musique électronique au début du XX^{ème} siècle. Ceux-ci étaient censés donner un accès instantané au jeu musical, y compris pour des novices en musique. Cependant, l’histoire du XX^{ème} siècle nous a appris que seule une poignée de musiciens ont réellement été reconnus dans leur pratique d’instruments électroniques tels que le Theremin et l’Onde Martenot. D’autre part, cette « intuitivité » est difficilement mesurable, et l’utilisation de systèmes commerciaux prônant une utilisation « intuitive » se montre rarement plus satisfaisante, ni même plus efficace.

D’après la littérature en psychologie et en design, il s’avère que la création et la pratique d’activités complexes, comme les arts ou les sports et leurs règles diverses et pointues, sont inhérentes à notre nature et à notre culture. Par exemple, le langage et l’écriture sont deux des activités les plus complexes qui soient, mais elles sont pourtant largement répandues et même nécessaires.

Depuis les années 50, de nombreux auteurs défendent l'idée que la simplicité d'utilisation d'un outil est en fait fonction de la complexité de la tâche à réaliser, et que la complexité n'amène pas de frustration si l'utilisation de l'outil est cohérente, compréhensible, et que la complexité de la tâche est justifiée par rapport à la complexité du résultat visé.

Les activités expertes sont complexes, mais elles représentent des combinaisons efficaces et éprouvées à travers le temps entre des tâches, des outils, des capacités perceptives et cognitives humaines, et des méthodes d'apprentissage efficaces.

Une partie de notre travail consiste à concevoir des systèmes interactifs inspirés par des activités expertes dont les non-experts ont une connaissance implicite, pour profiter de la puissance des activités expertes, tout en proposant des techniques d'interaction accessibles sans apprentissage poussé.

1. Sources d'inspiration en philosophie et en psychologie

L'étude de deux théories développées respectivement par la philosophie et la psychologie du XX^{ème} siècle nous permet d'appuyer notre argumentation, de définir l'expertise, et de décrire les capacités développées par les non-experts et la façon dont ils les acquièrent. Ces théories proposent des descriptions avancées des activités « médiées » par des outils (comme c'est le cas de la musique et de l'utilisation des systèmes informatiques), nous aident à comprendre comment et pourquoi un outil peut être accessible à un novice, et comment une connaissance déjà acquise peut être réutilisée dans une nouvelle situation, comme nous voulons le faire en basant des techniques d'interaction sur une connaissance implicite déjà acquise. Finalement, elles nous fournissent des directions pour estimer l'utilisabilité d'un système interactif autrement qu'en mesurant le temps requis pour réaliser une tâche élémentaire.

Tout d'abord, la phénoménologie décrit la perception comme une activité de compréhension, plutôt que comme la contemplation d'un phénomène existant a priori, comme souvent défini par certaines branches des sciences cognitives. Le sens que nous attribuons à notre environnement, à nos outils, et la réalité que nous percevons est construit par les phénomènes perceptifs, qui constituent la première interaction avec notre environnement. L'expertise est alors caractérisée par la capacité à structurer notre perception de différentes situations et la connaissance qui y est associée. Là où un novice peut manquer de clefs pour comprendre une situation et savoir comment atteindre ses objectifs, un expert sera capable d'identifier instantanément les aspects les plus importants et de savoir comment agir pour atteindre ses buts. La connaissance est forgée au cours des expériences successives, et dépend de nos capacités de compréhension.

Concernant l'interaction homme-machine, la phénoménologie nous indique qu'il est nécessaire de concevoir des abstractions pour aider les utilisateurs à faire sens des systèmes interactifs, et de rendre leur comportement compréhensible en donnant à voir leurs réactions et en guidant l'utilisateur pendant l'interaction pour l'aider à structurer sa perception du fonctionnement

de l'outil informatique. Deux mécanismes bien connus en IHM, appelés « feedforward » et « feedback », visent respectivement à montrer à l'utilisateur ce qu'il peut faire et comment le faire, et les actions que le système a pris en compte ainsi que ses réactions. Pour être efficace, un système interactif doit permettre à l'utilisateur de se concentrer sur ses objectifs plutôt que de passer du temps à comprendre son fonctionnement s'il n'est pas cohérent, ou si sa cohérence n'est pas lisible.

Deuxièmement, la théorie de l'activité se concentre sur les activités médiées par des outils, décrit l'expertise liée à l'utilisation d'outils, et « l'internalisation » des activités. L'élément fondamental de l'activité est l'action, une tâche élémentaire que l'utilisateur réalise consciemment, et qui n'est pas forcément dirigée vers son but global. Lorsque les actions sont bien connues et que l'expertise se développe, elles deviennent des opérations, réalisées inconsciemment et automatiquement. D'après la théorie de l'activité, nous n'utilisons pas que des artefacts techniques (i.e. les outils physiques comme un marteau, une pince, une machine ou un instrument de musique), mais également des artefacts psychologiques (comme les règles mathématiques, les langages ou d'autres systèmes symboliques). Pour la musique, nous pouvons dire que le rythme, l'harmonie et la tonalité sont des artefacts psychologiques utilisés à la fois par les musiciens et par les auditeurs.

Ces deux sortes d'artefacts sont plus que de simples outils. Les utiliser oriente notre compréhension, et ils représentent d'une certaine manière l'activité à laquelle ils sont associés. Ainsi, les non-musiciens voient leur perception de la musique affectée par les artefacts psychologiques qu'ils ont construit implicitement en écoutant de la musique, et ont accès à une part de la connaissance experte de la musique. La théorie de l'activité décrit le processus d'internalisation, qui consiste à devenir capable d'imaginer des actions et leurs résultats sans avoir besoin de les réaliser physiquement. Plus la connaissance est pointue, plus les phénomènes internalisés sont précis.

Lorsqu'un utilisateur sait exactement quels résultats il peut obtenir avec un outil, quels objectifs cela lui permettra d'atteindre, comment les atteindre et comment manipuler cet outil, l'outil devient un « organe fonctionnel » ou « instrument », caractéristique d'une utilisation experte. Pour permettre le développement de cette expertise, un outil doit montrer comment il doit être manipulé, et comment il doit être utilisé étant donné les buts de l'utilisateur. Pour l'utilisateur, un instrument est associé à des schémas d'utilisation, et l'utilisateur pourra les accommoder pour adapter sa connaissance à des nouvelles situations et de nouvelles tâches, ou assimiler ces dernières aux schémas existants. Ce concept permet d'expliquer comment des utilisateurs, novices ou experts, peuvent réutiliser leur connaissance dans diverses situations, comme nous voulons le faire en concevant des systèmes interactifs inspirés par la pratique musicale et réutilisant la connaissance musicale des non-musiciens.

En IHM, des approches comme l'interaction instrumentale de Beaudouin-Lafon ou le modèle humain-artefacts de Bødker et Klokmoose, présentés dans le manuscrit, tirent parti de ces théories pour définir l'utilisation de systèmes interactifs comme l'utilisation d'un instrument. D'après l'interaction

instrumentale, trois processus et caractéristiques complémentaires sont nécessaires à la standardisation et à l'efficacité d'un instrument : la réification (le fait de créer un instrument pour répondre à un besoin donné), le polymorphisme (le fait qu'un instrument puisse s'adapter à diverses situations), et la réutilisation (le fait de pouvoir réutiliser le résultat d'une utilisation précédente de l'instrument). Pour valider des modèles d'interaction comme l'interaction instrumentale, Beaudouin-Lafon et Mackay proposent d'en étudier les pouvoirs descriptifs, évaluatifs et génératifs. En effet, un modèle d'interaction pourra être plus largement utilisé et sera plus pertinent s'il permet de décrire et d'évaluer l'existant, mais aussi de créer des designs innovants. Mais l'interaction instrumentale se concentre sur l'efficacité d'un outil pour réaliser une tâche donnée, et étudie donc la relation entre l'outil et la tâche. Notre approche se concentre plutôt sur le lien entre les capacités de l'utilisateur et le fonctionnement de l'outil, que nous nommons utilisabilité.

De nombreux systèmes interactifs conçus récemment ne considèrent pas le développement de l'expertise, tendent à morceler l'interaction en une multitude de tâches élémentaires souvent difficiles d'accès plutôt que de permettre à l'utilisateur de s'investir pour internaliser la tâche et développer des automatismes. Même les interactions dites « naturelles », comme l'interaction gestuelle, nécessitent finalement que l'utilisateur apprenne comment fonctionnent les algorithmes de reconnaissance de gestes, et découvre par lui-même le vocabulaire d'actions disponible. Si des courants de la conception d'interactions, comme l'interaction tangible, réutilisent des objets du quotidien pour contrôler les ordinateurs, peu d'intérêt a été accordé à la réutilisation d'artefacts psychologiques, et à l'étude de la cohérence entre le fonctionnement d'une technique d'interaction et les capacités des utilisateurs.

2. L'expertise des non-experts

Dans la suite du manuscrit, nous présentons des définitions plus concrètes de l'expertise qui valident les observations fournies par la phénoménologie et la théorie de l'activité, et qui expliquent pourquoi et comment des non-experts peuvent avoir une connaissance avancée de certaines pratiques expertes. Nous caractérisons ensuite l'expertise musicale, pour finalement identifier les capacités avancées développées par les non-musiciens lorsqu'ils écoutent de la musique ou assistent à des concerts.

D'après la littérature sur l'expertise, la « pratique délibérée » est nécessaire pour développer de la connaissance et acquérir des compétences. Ainsi, les experts développent une meilleure organisation de leur connaissance, obtiennent de meilleures performances sur certaines tâches (en termes de qualité du résultat et en termes de rapidité), analysent les problèmes de façon plus efficace, et sont capables de planifier leurs actions et de détecter d'éventuelles erreurs. En plus de ces capacités cognitives, ils développent également des capacités physiques particulières. Par exemple, un sculpteur va développer des gestes dont la précision spatiale et temporelle sont inaccessibles à un novice. Ils sont aussi capables de combiner des actions élémentaires en actions plus complexes.

Cependant, dans l'étude des interactions sociales, plusieurs chercheurs ont montré que la connaissance est accessible par l'observation, sans forcément nécessiter une pratique délibérée personnelle. Ces théories permettent d'expliquer comment les jeunes enfants apprennent le langage sans en connaître les règles de façon explicite et exhaustive. Lors de la communication, qu'elle soit par exemple verbale ou artistique, une connaissance commune se construit entre ceux qui communiquent, qui va au-delà des connaissances personnelles respectives. Dans les sociétés actuelles, cette façon de créer et de partager de la connaissance est même fondamentale, puisque le sens que nous attribuons à notre environnement dépend largement du consensus global. Ces interactions sociales sont le lieu de la création d'une connaissance implicite importante. Par exemples, des musiciens et leur public construisent ensemble le résultat musical lors des concerts. Les musiciens produisent de la musique, et le public y réagit en manifestant sa satisfaction ou son mécontentement, qui influencent à leur tour les musiciens.

Cette faculté à comprendre les actions observées de façon implicite et à construire une connaissance commune est appelée intersubjectivité. D'autres processus, tels que notre capacité innée à imiter les actions observées, nous permettent de développer des compétences particulières grâce aux simples mécanismes perceptifs, par exemple visuels ou auditifs. La récente découverte des neurones miroirs vient appuyer ces observations, en montrant que certains neurones simulent les actions observées, pour en permettre une compréhension plus profonde, et préparer une éventuelle imitation. Comme ces processus de compréhension ne nécessitent pas de pratique délibérée particulière, nous pouvons dire qu'ils sont parmi les plus innés et les plus simples que nous sommes capables de mettre en œuvre, et assurément un domaine dans lequel nous sommes tous experts.

De nombreuses études ont été menées pour identifier et décrire les capacités expertes des musiciens. La connaissance musicale est composée de nombreuses règles comme le rythme, la tonalité ou l'harmonie, que la plupart des musiciens apprennent de façon explicite. Au cours de leur apprentissage, les instrumentistes développent des compétences physiques, leur permettant par exemple de synchroniser, ou au contraire de dissocier les mouvements de leurs membres ou de leurs doigts. De nombreuses méthodes d'apprentissage incitent les instrumentistes à d'abord morceler la complexité du geste instrumental, pour en apprendre séparément les différentes composantes. Les capacités cognitives sont également nombreuses, permettant aux instrumentistes experts de planifier leurs actions, d'anticiper leur résultat sonore, ou encore de percevoir rapidement leurs erreurs et de pouvoir les corriger. Ils développent ainsi une faculté d'expression à travers la musique, en utilisant les règles élémentaires de la musique, en les combinant, et en y rajoutant des accentuations ou des modulations rythmiques personnelles pour souligner leur interprétation.

Ces processus définissent une syntaxe, une grammaire, et une expression musicale qui sont connues de façon implicite par les non-musiciens. Nombreux sont ceux qui ont des goûts musicaux marqués, ou qui sont capables de distinguer différentes interprétations d'un même morceau. La première raison à cette sensibilité globale à la musique est son lien avec notre fonctionnement physique : plusieurs organes fonctionnent de façon périodique, nous marchons

et nous courons en rythme, les variations temporelles de la musique correspondent à l'évolution temporelle de nos mouvements, et la hauteur et les intonations de la voix sont les deux premiers indices qu'ont les jeunes enfants pour comprendre l'expression des adultes. Ensuite, nous comprenons la structure musicale grâce à la régularité des événements, et la récurrence des structures rythmiques, harmoniques et tonales. Nos attentes deviennent si fortes que notre cerveau réagit de la même manière si une phrase musicale ne respecte pas les règles de la musique occidentale tonale que si une phrase verbalisée ne respecte pas la syntaxe ou la grammaire du langage. Certains auteurs insistent même sur le fait que la connaissance musicale des musiciens et celle des non-musiciens présentent plus de similitudes que de différences.

D'autre part, nous sommes tous habitués à siffler, à fredonner, à danser, ou à taper des rythmes simples en écoutant des morceaux de musique. Dès le plus jeune âge, les jeunes enfants sont capables de reproduire une mélodie qu'ils viennent d'entendre, et d'en produire de nouvelles. La perception visuelle est également importante : les gestes du musicien et ses expressions faciales viennent souligner certaines composantes de la structure musicale ou de l'interprétation. Mais les pratiques courantes en musique électronique réduisent cette cohérence visuelle, et appauvrissent donc la compréhension du public.

3. Réutiliser la connaissance musicale implicite en IHM

Nous avons identifié les utilisateurs novices comme experts, en tant qu'êtres humains, en compréhension incarnée, en médiation de leurs activités par des artefacts techniques et psychologiques, en compréhension par l'observation et en imitation.

Notre objectif est de créer des systèmes interactifs basés sur ces capacités innées, et sur la connaissance qu'elles ont permis d'acquérir au préalable. De plus, la musique est une activité qui apporte de la satisfaction, dont des techniques d'interaction pourraient tirer parti pour inciter les utilisateurs à s'y investir. Contrairement à la reconnaissance vocale qui tire parti de notre connaissance du langage, des techniques d'interaction basées sur notre connaissance de la musique auraient moins de chances de frustrer les utilisateurs, car nos attentes, bien que conséquentes, ne sont pas aussi strictes. De plus, nous observons que dans des jeux vidéo comme Guitar Hero, des non-musiciens sont capables d'atteindre rapidement un niveau d'expertise élevé, grâce à des mécanismes de feedback et de feedforward appropriés.

A partir de ces observations, nous proposons un cadre de conception pour définir et étudier l'utilisabilité et l'expressivité des instruments d'interaction. Nous définissons l'utilisabilité comme répondant à quatre critères : la compréhensibilité, la simplicité de manipulation, la simplicité d'apprentissage, et le potentiel d'attrance.

Pour être facilement compréhensible, un instrument doit avoir un fonctionnement cohérent, rendre ses réactions lisibles, et abstraire le fonctionnement de l'ordinateur et des applications. D'après les études citées précédemment, un instrument aura plus de chances d'être compréhensible s'il

met l'utilisateur face à des situations qu'il connaît déjà, voire si son fonctionnement correspond à de la connaissance qui a déjà été internalisée. C'est le cas de la reconnaissance vocale, de l'interaction tangible, des simulations physiques ou de l'utilisation de stylos interactifs. D'après l'étude de l'intersubjectivité, la compréhensibilité pourra également être favorisée si le fonctionnement de l'instrument correspond à une activité qui a déjà été observée par l'utilisateur, comme c'est le cas de Guitar Hero.

Un instrument sera facilement manipulable s'il est utilisé en effectuant des gestes naturels, comme l'interaction par le regard, ou s'il utilise des gestes qui ont été automatisés, comme l'interaction tangible avec des objets du quotidien.

Si le fonctionnement d'un instrument n'est pas immédiatement compréhensible, ou si celui-ci n'est pas immédiatement manipulable, son apprentissage doit être facilité. Il existe différentes méthodes d'apprentissage que l'on peut trouver dans la littérature en IHM : les guides dynamiques, les systèmes de démonstration dans le contexte applicatif ou hors contexte (par exemple des vidéos pour apprendre des gestes), ou des systèmes d'apprentissage par images (par exemple une photographie d'une main en train de faire un geste). Les guides dynamiques guident l'utilisateur pendant qu'il fait le geste, en fournissant des mécanismes de feedback et de feedforward, et laissent l'utilisateur découvrir la complexité du système de façon progressive, ce qui favorise l'apprentissage d'après la phénoménologie. En IHM, diverses études se sont intéressées aux composantes nécessaires pour encourager les utilisateurs à développer une expertise : ce qu'ils apprennent dès le début doit être réutilisé quand ils deviennent experts, le fait d'utiliser une nouvelle technique d'interaction ne doit pas changer radicalement leurs tâches, et les résultats de leurs actions doivent pouvoir être prévus.

Nous définissons différents types d'expressivité pour les instruments d'interaction : la largeur sémantique (le nombre d'actions qui sont prises en compte), la variété sémantique (la richesse des différences entre les actions prises en compte), la largeur syntaxique (le nombre de règles syntaxiques selon lesquelles les actions peuvent être combinées) et la variété syntaxique (la richesse des différences entre ces règles syntaxiques).

Le potentiel d'attraction d'un instrument d'interaction est défini par son accessibilité initiale, l'intérêt initial qu'il suscite chez l'utilisateur, et la satisfaction d'utilisation.

Notre proposition de nous inspirer de pratiques expertes dont les non-experts ont une certaine connaissance permet de favoriser l'expressivité (car c'est une caractéristique propre aux activités expertes), la compréhensibilité (car l'instrument aura un fonctionnement que l'utilisateur est déjà capable de comprendre), et la simplicité d'apprentissage (car les activités expertes ont des méthodes d'apprentissage efficaces). D'autre part, des techniques d'interaction basées sur la musique peuvent être plus attractives car elles rappelleront à l'utilisateur une activité qui lui procure de la satisfaction.

4. Utiliser des motifs rythmiques pour l'interaction

Dans le premier projet présenté dans le manuscrit, nous utilisons les capacités naturelles des non-musiciens pour percevoir, comprendre et reproduire des structures rythmiques. Nous étudions la faisabilité de l'utilisation de motifs rythmiques pour l'interaction, nous proposons des façons de concevoir des systèmes basés sur le rythme, et nous proposons des solutions techniques dont nous démontrons l'efficacité pour en encourager l'utilisation.

Ce projet réutilise la majeure partie de nos observations sur la connaissance implicite et sur l'intersubjectivité. Dans le cadre de conception présenté précédemment, notre technique d'interaction utilisant des motifs rythmiques peut être définie comme compréhensible (car les utilisateurs ont une bonne connaissance du rythme), et simple à manipuler (car les utilisateurs sont habitués à bouger en rythme et à reproduire des rythmes simples). Ainsi, nous ne nous intéressons à la simplicité d'apprentissage que dans la mesure où certains feedback (auditif et/ou visuel) peuvent faciliter la reproduction et la mémorisation des motifs rythmiques.

L'utilisation de motifs rythmiques pour l'interaction présente d'autres avantages. Par exemple, le fait que les motifs rythmiques puissent être réalisés avec de nombreux dispositifs d'entrée (claviers, souris, écran tactile...), ou le fait que l'interaction rythmique soit plus adaptée aux dispositifs mobiles que l'interaction gestuelle car les motifs rythmiques ne nécessitent qu'une petite surface d'écran.

Nous définissons plusieurs règles pour créer des motifs rythmiques suffisamment expressifs (en termes de largeur sémantique et de variété sémantique) sans pour autant amener la même complexité que les rythmes musicaux. Nous proposons trois types de « taps »¹ (impulsion, un temps et deux temps), et trois types de « pauses » (relâchement entre deux taps consécutifs, un temps et deux temps) (voir fig. 1).

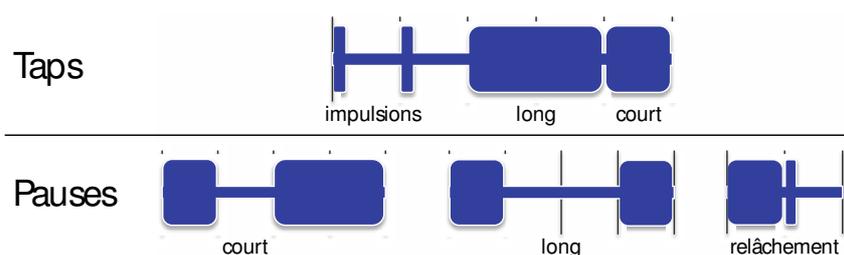


Figure 1 : les types de taps et de pauses pour définir nos motifs rythmiques.

Dans nos expérimentations, nous choisissons un tempo de 120 pulsations par minutes, qui est un tempo standard en musique et une valeur commune pour diverses actions motrices.

¹ Nous utilisons le mot « tap » pour les événements rythmiques, car dans nos expérimentations, les utilisateurs tapent sur un pavé tactile avec le bout de l'index pour réaliser les motifs rythmiques

Dans une première expérimentation, nous demandons à 12 utilisateurs de reproduire des motifs rythmiques. Le stimulus est auditif et visuel (ils voient la forme du motif rythmique se remplir à l'écran, et il entendent un son pour chaque tap). Nous testons quatre conditions de feedback lorsque l'utilisateur reproduit le motif rythmique : visuel, audio, audiovisuel et une condition sans aucun feedback. Nous testons 30 motifs rythmiques qui font entre deux et six temps, et ont entre deux et six taps.

Avec un reconnaiseur strict conçu spécialement pour tester la précision de reproduction lors de cette expérimentation, nous observons que les utilisateurs reproduisent les motifs rythmiques moins précisément quand ils n'ont pas de feedback, alors que les deux conditions qui ont les meilleurs taux de réussite moyens sont les conditions avec les feedback audio et audiovisuel. 15 motifs rythmiques sont reconnus plus de 70% du temps par le reconnaiseur strict, ce qui montre que des utilisateurs novices sont capables d'une grande précision sur un vocabulaire déjà relativement expressif. Nous observons également que plus le motif comporte de taps ou plus il est long, moins les résultats de reproduction sont bons. La moitié des participants préfèrent le feedback audio aux autres, et un seul utilisateur a dit éprouver des difficultés à reproduire les motifs rythmiques. En concevant un classifieur moins strict et plus adapté à une utilisation applicative, les résultats de cette expérimentation atteignent 94% de réussite en moyenne pour l'ensemble des utilisateurs, et pour chaque nombre de taps ainsi que chaque nombre de temps.

Dans une deuxième expérimentation divisée en deux parties présentées lors de deux journées consécutives, nous étudions la mémorisation des motifs rythmiques avec seulement un feedback audio lorsqu'ils sont utilisés pour déclencher des commandes représentées par 14 images montrant des objets ou des fruits (voir Fig. 2). Ces résultats sont comparés aux résultats de mémorisation obtenus dans les mêmes conditions et pour les mêmes commandes avec des raccourcis clavier. Chacun des 14 utilisateurs passe d'abord par une courte phase d'apprentissage où chaque motif rythmique lui est présenté avec le nom de la commande associée, et doit être reproduit. Vient ensuite une phase de restitution où un nom de commande lui est présenté et il doit réaliser le motif rythmique associé, s'il s'en souvient. Sinon, il est invité à appeler un mécanisme d'aide qui lui montre à nouveau le motif rythmique associé à cette commande. Le lendemain, les participants passent une autre phase de restitution, cette fois et sans apprentissage. A la fin de l'expérimentation, les participants ont le choix, pour chaque commande, entre réaliser le raccourci clavier et réaliser le motif rythmique. Les résultats montrent que les motifs rythmiques obtiennent d'aussi bons résultats que les raccourcis clavier qui sont pourtant une des techniques d'interaction les plus communes, que le taux de mémorisation moyen à la fin de la session de la première journée (environ 40 minutes) est supérieur à 90% pour chaque technique, et comparable le lendemain. Le taux d'appel à l'aide chute à partir de la première journée, pour passer en dessous de 5%. Dans la partie finale de la deuxième journée, 10 participants ont utilisé les motifs rythmiques plus souvent que les raccourcis clavier, et sept d'entre eux ont utilisés les motifs rythmiques plus de 80% du temps, ce qui montre que la technique d'interaction a plu aux utilisateurs.

CMD1	CMD2	CMD3	CMD4	CMD5	CMD6	CMD7
						
R1 = P20	R2 = P11	R3 = P10	R4 = P9	R5 = P19	R6 = P4	R7 = P3
						
Ctrl+Y	Shift+H	Ctrl+X	Shift+E	Ctrl+R	Shift+F	Ctrl+N
CMD8	CMD9	CMD10	CMD11	CMD12	CMD13	CMD14
						
R8 = P2	R9 = P1	R10 = P29	R11 = P18	R12 = P6	R13 = P28	R14 = P12
						
Shift+B	Ctrl+D	Shift+T	Ctrl+H	Shift+G	Ctrl+A	Shift+W

Figure 2 : le vocabulaire utilisé pour l'expérimentation de mémorisation des motifs rythmiques

D'autre part, pour certains motifs rythmiques, des utilisateurs ont construit par eux-mêmes des moyens mnémotechniques pour les mémoriser, alors qu'aucun participant n'a eu recours à une telle stratégie de mémorisation pour les raccourcis clavier.

Finalement, nous avons proposé une nouvelle modalité d'entrée qui s'est avérée simple à manipuler, compréhensible, simple à apprendre avec uniquement un feedback audio, attirante, et facile à internaliser car les utilisateurs ont développés des moyens mnémotechniques. D'autre part, nous avons proposé des règles pour définir des vocabulaires de motifs rythmiques expressifs et un reconnaissanceur utilisable dans un contexte applicatif.

5. Arpege : Conception et apprentissage de postures multi-doigts

Dans ce deuxième projet, nous étudions la conception et l'apprentissage de postures multi-doigt sur des écrans multi-tactiles (voir Fig. 3). Bien que les non-musiciens ne soient pas habitués à reproduire et apprendre des vocabulaires de postures multi-doigts (contrairement aux pianistes ou aux guitaristes par exemple), nous présentons des études qui soutiennent que des non-musiciens peuvent facilement reproduire quelques accords de piano pour accompagner un morceau de musique. Nous pouvons donc dire que dans le cas où les positions de chaque doigt sont définies et le nombre de combinaisons possibles est limité, les vocabulaires de postures multi-doigts devraient être accessibles aux novices.

D'autre part, comme nous l'avons dit précédemment, créer des méthodes d'apprentissage inspirées par la pratique des instruments de musique peut permettre aux utilisateurs de progresser plus rapidement dans leur apprentissage du système. Une telle méthode d'apprentissage pourrait palier au fait que des techniques d'interaction basées sur des vocabulaires de postures multi-doigts sont limitées du point de vue de la compréhensibilité et de la simplicité de manipulation.

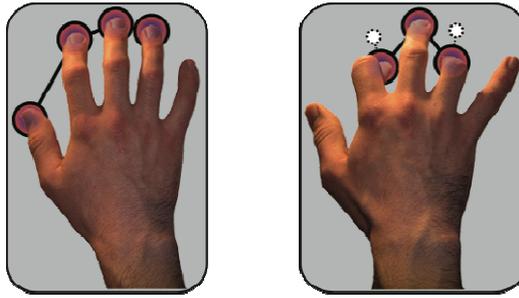


Figure 3 : Deux exemples de postures multi-doigts

Tout d'abord, nous analysons des études sur les capacités motrices et les contraintes mécaniques de la main, pour définir des vocabulaires de postures multi-doigts qui soient à la fois expressifs et utilisables. En plus des combinaisons obtenues à partir de la position « relâchée » de la main (i.e. en posant le bout des doigts sur l'écran tactile en évitant les tensions musculaires dans chaque doigt, pour que chaque doigt soit dans la position qui semble la plus naturelle à l'utilisateur), et après avoir éliminé les cinq configurations dans lesquelles soit le majeur soit l'annulaire sont levés de l'écran tactile alors que leurs voisins sont en contact avec l'écran (que nous identifions comme les configurations les plus inconfortables), nous définissons des positions supplémentaires pour chaque doigt (voir Fig.4). Nous appelons ces nouvelles positions de chaque doigt les positions « tendues », et obtenons un vocabulaire de 480 postures multi-doigts en ajoutant les postures incluant des doigts en position « tendue ».

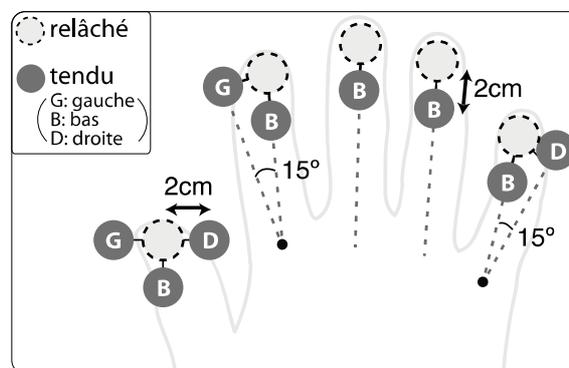


Figure 4 : Positions relâchées et tendues pour chaque doigt.

Dans une première expérimentation, nous demandons à 12 utilisateurs de reproduire un vocabulaire de 52 postures multi-doigts (les 26 où les doigts ne sont que dans leurs positions « relâchées » et un vocabulaire représentatif de 26 postures « tendues ») et d'évaluer la compréhensibilité et le confort d'utilisation pour chaque posture. Les résultats montrent que les postures « tendues » sont perçues comme moins confortables que les autres, que plus de doigts sont impliqués, moins la posture est compréhensible et confortable d'après les participants, et que les configurations que nous avons identifiées comme étant potentiellement les moins confortables ont été jugées comme telles par les participants uniquement dans le cas des postures « relâchées ».

Nous présentons ensuite la conception d'un guide dynamique pour apprendre des postures multi-doigts associées à des noms de commandes, inspiré par l'apprentissage des accords en musique (voir Fig.5). L'utilisateur

pose les doigts un à un sur des cibles affichées par le guide, en suivant les informations de positionnement données par les labels des commandes (les flèches à côté du nom de la commande indiquent la position adéquate). Une commande est déclenchée quand tous les doigts sont levés de l'écran simultanément alors que la posture correcte avait été atteinte.

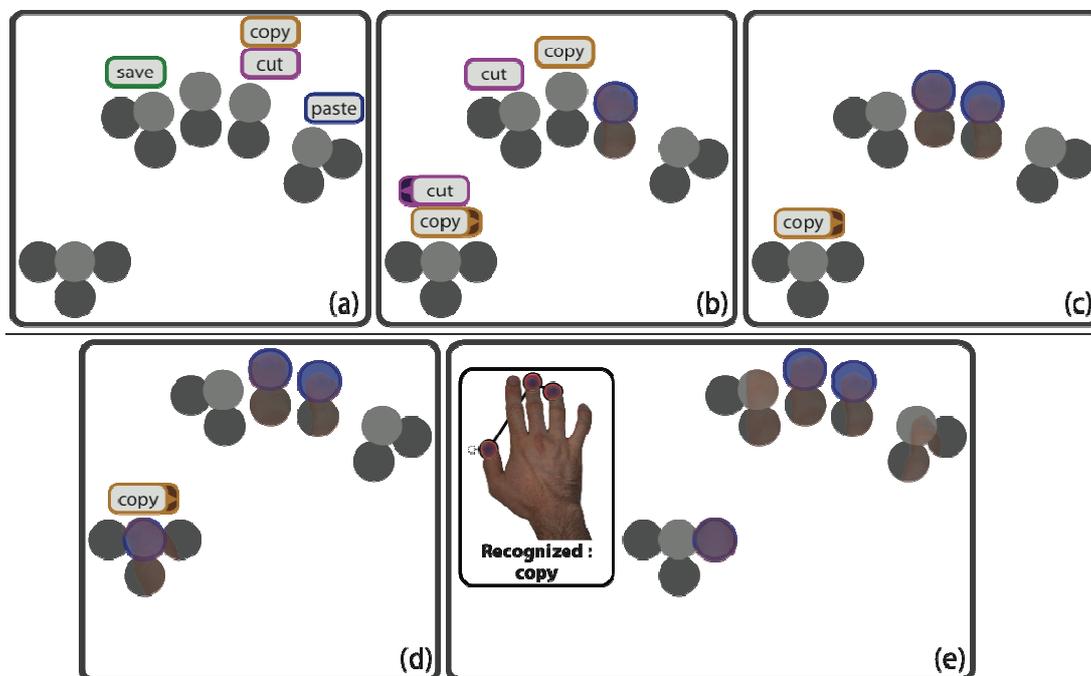


Figure 5 : Le guide dynamique Arpège pour postures multi-doigts guide les utilisateurs pour qu'ils posent les doigts les uns après les autres afin d'apprendre une posture associée à une commande. Ici l'utilisateur veut déclencher la commande « copy ».

Dans une seconde expérimentation, nous comparons l'apprentissage des postures avec notre guide dynamique, Arpège, à l'apprentissage avec des images montrant une main en train de réaliser la posture. Un groupe de 12 participants apprend les 12 postures avec Arpège, alors qu'un autre groupe de 12 les apprend avec les images. Comme pour le projet précédent, nous testons la mémorisation sur deux jours. Les résultats ne montrent pas de différence significative pour le taux de mémorisation ni pour le taux d'appel à l'aide entre les deux techniques, mais les participants novices sont capables de mémoriser 90% de nos postures multi-doigts à la fin de la session de la première journée.

D'autre part, la méthode Arpège présente certains avantages qui peuvent s'avérer cruciaux dans certains contextes applicatifs : elle aide les utilisateurs à créer des moyens mnémotechniques, elle présente la complexité des accords de façon progressive, elle assure que l'utilisateur a réalisé la posture de façon adéquate (alors que cette vérification n'est pas possible dans le cas de l'apprentissage avec les images), et elle permet à l'utilisateur de corriger ses erreurs éventuelles et d'explorer le vocabulaire.

6. Une stratégie de mapping pour des logiciels de musique « instrumentaux »

Avec l'essor des outils informatiques de création musicale, le jeu musical a perdu une part de sa matérialité, et de ses cohérences causales et énergétiques entre le geste et le résultat sonore. Par conséquent, les concerts sont moins compréhensibles pour le public, pour qui il est difficile de savoir quel geste a déclenché quel son. De plus, cette perte de cohérence rend les systèmes informatiques moins accessibles que les instruments acoustiques (il est plus difficile de jouer quelques sons avec un logiciel de musique inconnu qu'avec la plupart des instruments de musique même si l'on n'en a jamais joué), et la pratique moins stable. En effet, les instruments acoustiques permettent aux musiciens de reproduire précisément des structures musicales et d'y introduire des variations, pour progressivement apprendre à jouer, ce qui n'est pas toujours possible avec les systèmes informatiques.

En nous appuyant sur diverses études des instruments acoustiques, nous en proposons une décomposition fonctionnelle centrée sur les mécanismes de conversion d'énergie (voir Fig. 6). En effet, tout en maintenant la cohérence énergétique entre le geste du musicien et le résultat sonore, un instrument de musique donne accès à des mécanismes physiques vibratoires irréalisables avec le corps humain, et cette conversion d'énergie constitue l'essence de ses propriétés fonctionnelles.

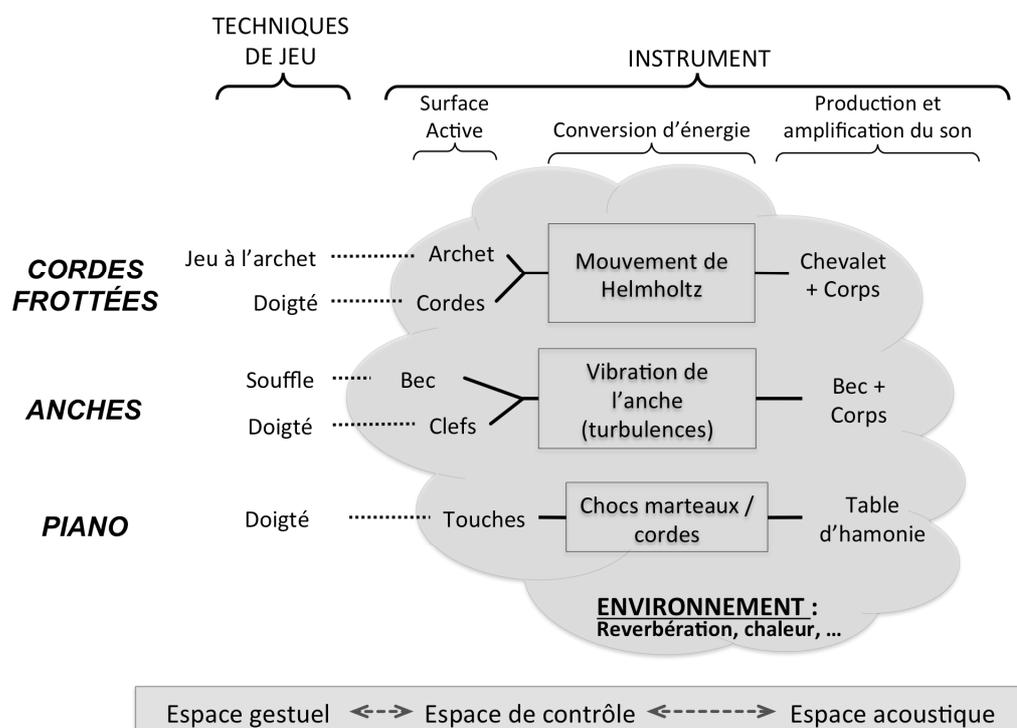


Figure 6 : Décomposition fonctionnelle des instruments acoustiques, centrée sur les mécanismes de conversion d'énergie.

Pour les logiciels de jeu musical, les problématiques de mapping de paramètres (le lien entre les paramètres des dispositifs d'entrée, et les paramètres des moteurs de synthèse sonore) ont souvent été décrites comme le point crucial de leur « instrumentalité ». Nous observons également d'autres concepts qui visent à améliorer la « matérialité » des systèmes informatiques, comme les métaphores ou la métonymie.

Parmi les stratégies de mapping explorées dans la littérature, certaines mettent l'utilisateur face à des abstractions ou des systèmes dynamiques dont il peut appréhender le comportement synthèse (comme les espaces d'interpolation de paramètres, les modèles statistiques ou les modèles physiques) plutôt que de contrôler de façon exhaustive l'ensemble des paramètres du moteur de.

Nous proposons un cadre de conception pour le « mapping par modèles de comportement », qui met les mécanismes de conversion d'information et les abstractions au centre du système (voir Fig. 7.). Un modèle de comportement est un système dynamique ayant une représentation graphique jouant le rôle d'abstraction, qui reçoit les paramètres des dispositifs d'entrée et contrôle à son tour le moteur de synthèse. L'utilisateur a accès à certains de ses paramètres (ou « poignées »). Des mesures de son évolution (ou « mesures comportementales ») sont alors envoyées au moteur de synthèse. Ce cadre de conception peut être utilisé pour décrire les stratégies de mapping existantes. Nous proposons également des directives basées sur la littérature pour créer des modèles de comportement qui rapprocheraient l'utilisation de logiciels de musique de l'expérience du jeu instrumental en musique.

Parmi les directives que nous proposons, certaines tendent à rendre le jeu musical avec des logiciels de musique plus accessibles aux utilisateurs novices, d'autres à permettre à des utilisateurs experts de développer une virtuosité à part entière.

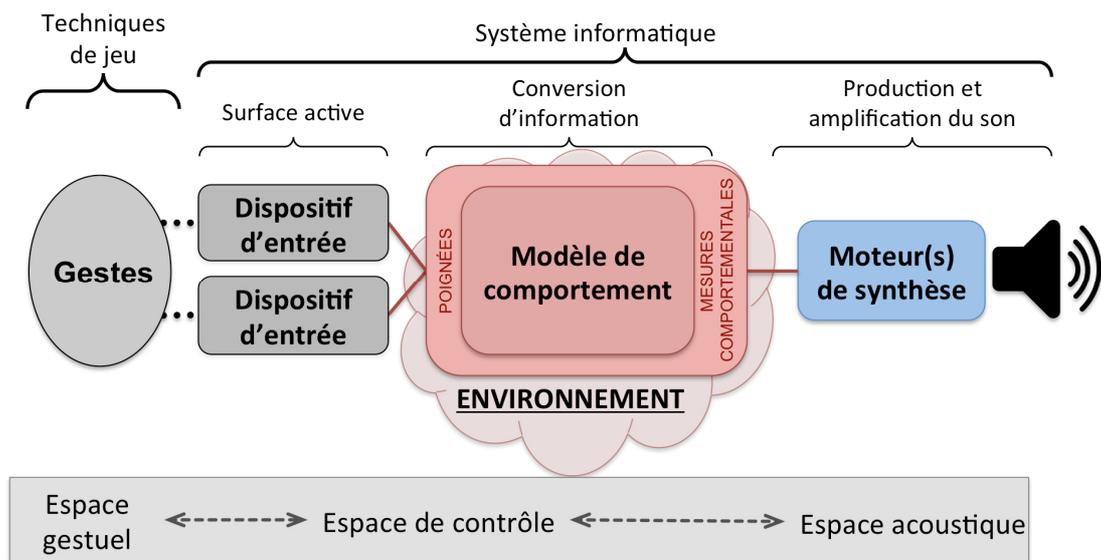


Figure 7 : Le mapping par modèles de comportement

Nous présentons également une architecture logicielle qui simplifie à de nombreux égards l'implémentation et l'évaluation des stratégies de mapping et des modèles de comportement. En facilitant les combinaisons entre dispositifs d'entrée, modèles de comportement et moteurs de synthèse, notre architecture logicielle rapproche la conception logicielle des procédés de lutherie acoustique, où les facteurs d'instruments ont passé des siècles à combiner des méthodes d'interaction avec des mécanismes de conversion d'énergie et des mécanismes de production sonore pour créer les instruments présents aujourd'hui dans l'orchestre. De la même manière, une taxonomie des modèles de comportement est aujourd'hui envisagée, et nous pourrions la mettre en parallèle des taxonomies de dispositifs d'entrée et de moteurs de synthèse sonores existants dans la littérature pour explorer l'ensemble des combinaisons possibles, et découvrir les combinaisons les plus « instrumentales ».

De par sa modularité, notre architecture permet également de profiter des possibilités supplémentaires accessibles grâce aux moyens informatiques, en permettant notamment de combiner des modèles de comportement pour créer des comportements plus complexes, comme nous le présentons dans le manuscrit.

Conclusion

Avec ces trois projets, nous montrons que, dans ces cas :

- la connaissance musicale implicite des non-musiciens peut être réutilisée en interaction ;
- les méthodes d'apprentissage expertes sont efficaces et adaptées à nos facultés de compréhension ;
- le fait de prendre les instruments acoustiques comme source d'inspiration peut aider à concevoir des logiciels de musique expressifs et utilisables.

Nous proposons l'étude de l'utilisabilité comme une alternative à l'immédiateté recherchée par les systèmes interactifs commerciaux, et nous proposons des méthodes pour tirer parti de la richesse des pratiques expertes et de la connaissance implicite des non-experts pour créer des systèmes interactifs utilisables et expressifs.