

State of the art
Whisperer

December 17, 2021

Étudiants :

Marion	CANDERLE	marion.canderle1@etu.univ-lorraine.fr
Anaël	ANGELINI	anael.angelini9@etu.univ-lorraine.fr
Pénélope	DEFOURNE	penelope.defourne7@etu.univ-lorraine.fr

Encadrants :

Stéphanie	FLECK
Stéphane	FAEDDA
Pierre	CARIO

Mots-clés: Interactions Humain-machine, expérience utilisateur, apprentissage, écoute active, collaboration, cognition incarnée, cognition multimodale.

Sommaire

Introduction	1
1 Attention	3
1.1 Attention sélective.....	3
1.2 Attention divisée	4
2 Mémoire de travail	7
2.1 Modèle de Baddeley et Hitch.....	7
2.2 Lien entre mémoire de travail et attention	8
3 Les théories de l'apprentissage	9
3.1 La théorie cognitive de l'apprentissage multimédia.....	9
3.2 La théorie de l'apprentissage multisensoriel.....	11
3.3 L'apprentissage collaboratif.....	12
4 Bilan	13
5 Conception du <i>Whisperer</i>	14
5.1 Conception centrée sur les utilisateurs	14
5.2 Forme du <i>Whisperer</i>	15
5.3 Caractéristiques du <i>Whisperer</i>	17
Conclusion.....	19
Bibliographie.....	20
Annexes.....	26
Annexe 1 - Premier prototype du <i>Whisperer</i> , forme de deux croches avec enceintes amovibles.....	26
Annexe 2 - Deuxième prototype <i>Whisperer</i> , forme de deux croches directement préhensibles.....	27

Introduction

La présente étude s'inscrit dans la continuité du projet e-TAC (Environnements Tangibles et Augmentés pour l'Apprentissage Collaboratif) qui a pour objectif de concevoir des environnements numériques afin de contribuer à l'apprentissage collaboratif chez des élèves âgés de 9 à 15 ans.



Figure 1 - Photographies de l'interface CARDS (Collaborative Activities based on the Real and the Digital Superposition). (Auteur : P.Lay – Professeur de Technologies associé au projet e-TAC - Collège J. de la Fontaine St Avold)

Dans le cadre de ce projet, une interface tangible et augmentée pour l'apprentissage, nommée CARDS, a été conçue par l'équipe chargée du projet (Voir Figure 1). CARDS est un système de réalité mixte et interactif dont l'objectif est de promouvoir l'apprentissage collaboratif à l'école par la manipulation physique de cartes augmentées. Développée par un processus de conception itératif, l'interface CARDS présente un bon potentiel pédagogique dans le cadre d'activités collaboratives, ce que ne permettent pas les systèmes informatiques traditionnels (Giraudeau et al., 2019).

Les interfaces tangibles sont des interfaces qui "augmentent le monde physique réel en couplant des informations numériques aux objets et environnements physiques de tous les jours" (Fishkin, 2004). Par exemple, dans le cas de l'interface CARDS, les objets peuvent correspondre à des cartes (Voir Figure 1). Le but est de permettre aux utilisateurs de "saisir littéralement les données avec leurs mains" (Shaer et Hornecker, 2009). Ces interfaces impliquent des événements d'entrée liés à la manipulation des objets ce qui modifie le système informatique. Ce dernier produira alors une réponse (i.e. un événement de retour) passant par différentes

modalités sensorielles et qui informe les utilisateurs sur les conséquences de leurs actions.

Cependant, la modalité de transmission d'informations par la voie sonore (e.g. consignes, explications dans une vidéo) est généralement peu prise en compte dans ce type d'interface. De plus, l'utilisation de ce type d'outils engendre un niveau sonore en classe qui n'est pas propice à l'apprentissage. En effet, l'interface a été imaginée pour être utilisée en parallèle par plusieurs groupes d'élèves (4-5 élèves par groupes) au sein d'une même classe. De ce fait, les pistes audios sont souvent lancées de manière asynchrone par tous les îlots de travail et peuvent perturber les différents groupes. En outre, l'Organisation Mondiale de la Santé recommande un niveau de fond sonore de 35 dB maximum dans les salles de classe. Néanmoins, il semble que ce niveau soit largement dépassé en contexte scolaire (Jamieson et al., 2004 ; Fadeyi et al., 2014) en dehors de l'utilisation de ce type d'interface.

Or, un environnement scolaire bruyant constitue une gêne pour les tâches impliquant le langage et notamment la compréhension de la parole (Shield et Dockrell, 2003 ; Jamieson et al., 2004). La présence de bruit dans l'environnement a des effets délétères sur les performances d'apprentissage des enfants, ceux-ci devant produire un effort supplémentaire pour écouter sélectivement un signal auditif (Downs & Crum, 1978). Par ailleurs, un environnement bruyant peut entraîner un effet Lombard (Zollinger et Brumm, 2011), obligeant les élèves et leur professeur à parler plus fort, ce qui augmente encore plus le niveau sonore dans les classes. En conséquence, l'attention des enfants et de l'enseignant est rudement sollicitée dans ce type d'environnement par la présence de nombreux bruits distrayeurs

Ainsi, notre objectif est de concevoir une interface tangible sonore appelée *Whisperer* qui viendra compléter l'interface CARDS. D'une part, elle permettrait de diffuser des sons (incluant du contenu verbal) en complément de vidéos à un ou deux élèves, tout en veillant à ne pas accroître le niveau sonore de la classe. Elle devra également favoriser l'apprentissage collaboratif des élèves et solliciter leur attention auditive. L'ajout d'un tel dispositif constituerait aussi un avantage pour les enseignants qui pourraient passer moins de temps à maintenir une ambiance calme en classe pour pouvoir se consacrer pleinement à l'apprentissage.

Diffuser du son par le biais d'un haut-parleur augmenterait le niveau sonore global dans la classe et serait potentiellement délétère pour l'apprentissage. Par ailleurs, l'utilisation d'un casque audio classique ne serait pas propice au travail collaboratif, car cela pourrait avoir tendance à isoler les élèves et donc à diminuer leurs interactions. Ainsi, ces propositions ne seraient pas viables pour atteindre notre objectif et résoudre les problèmes précédemment cités.

La conception du *Whisperer* suivra un modèle de design centré sur l'utilisateur, nécessitant de bien comprendre les besoins des individus qui l'utiliseront pour leur offrir une expérience d'utilisation optimale. Pour cela, la conception d'un tel dispositif implique de s'intéresser préalablement à certains concepts issus des

sciences cognitives, tels que l'attention, la mémoire de travail, ainsi que certaines théories de l'apprentissage pour avoir une meilleure compréhension du fonctionnement cognitif.

1 Attention

1.1 Attention sélective

L'attention sélective peut être définie comme la capacité à “*orienter sélectivement [son] attention sur des informations jugées pertinentes (au regard de la tâche en cours) plutôt que sur des informations jugées non pertinentes*” (Maquettiaux, 2017). Ainsi, l'attention sélective implique de traiter et de filtrer les nombreuses informations qui parviennent à nos sens. Différentes propositions théoriques ont été énoncées pour tenter de savoir lors de quelle étape de traitement ce filtre s'opère.

D'abord, le modèle du filtre précoce de Broadbent (1958) propose que toutes les informations sensorielles bénéficieraient d'un traitement superficiel basé sur les propriétés physiques des stimuli. Ces informations seraient stockées dans notre mémoire immédiate préattentive, ou mémoire sensorielle (Sperling, 1960), sous forme de représentations mentales durant des durées très brèves. Puis, il y aurait un traitement sémantique qui s'opérerait de façon sérielle, c'est-à-dire sur une représentation à la fois. Cette représentation est sélectionnée par le filtre de l'attention si elle est pertinente par rapport au but à atteindre. Par la suite, Deutsch et Deutsch (1963) ont proposé le modèle du filtre attentionnel tardif. Dans celui-ci, tous les stimuli sensoriels bénéficient d'un traitement sémantique. C'est après cette analyse approfondie des informations que se situerait le filtre attentionnel, lors de l'étape de sélection de la réponse. Enfin, Treisman (1960, 1964, 1969) a émis l'hypothèse du filtre attentionnel atténué. Selon cette hypothèse, toutes les informations peuvent potentiellement être traitées sémantiquement. Il y aurait un mécanisme d'atténuation des informations à ignorer qui réduirait leur qualité et leur intensité. Cependant, certaines informations ont un bas seuil d'activation (e.g. notre prénom). Ainsi, elles requièrent peu d'attention et peuvent être identifiées, même si elles ne sont pas pertinentes.

De nombreux travaux ont tenté de trancher entre ces différents modèles. Tout d'abord, les expériences de *shadowing* en situation d'écoute dichotique (Moray, 1959 ; Corteen et Wood, 1972). Dans celles-ci, les individus sont soumis à deux messages simultanés : un pertinent à répéter à voix haute et un second non pertinent qui est à ignorer. Par moments, des mots spécifiques sont diffusés dans le message à ignorer. Le but est de voir si les sujets les remarquent. Les résultats montrent que les participants parviennent à détecter ces informations bien qu'ils n'y portent pas attention. Cela vient rejeter le modèle de Broadbent (1958) puisqu'il

implique qu'il n'y a pas d'analyse sémantique des informations à ignorer. Ainsi, ces premiers résultats vont dans le sens des modèles du filtre attentionnel tardif de Deutsch et Deutsch (1963) et du filtre atténué de Treisman (1960, 1964, 1969). Cependant, des travaux plus récents ont suggéré l'idée d'un déplacement très rapide de notre attention vers le message à ignorer (Lachter, Forster et Ruthruff, 2004). Cela a conduit à reconsidérer la théorie de Broadbent puisqu'en l'absence de ces déplacements attentionnels vers des informations non pertinentes (e.g. en les bloquant expérimentalement), celles-ci ne sont pas traitées sémantiquement. Ainsi, ces déplacements expliqueraient pourquoi nous pouvons parfois détecter des informations qui ne sont pas pertinentes. Ils pourraient aussi avoir un lien avec notre capacité à traiter simultanément plusieurs informations.

Dans la conception du *Whisperer*, il sera important de prendre en compte les points suivants :

- Le contenu des informations sonores transmises par le *Whisperer* devra être pertinent par rapport au(x) but(s) à atteindre afin que les élèves portent préférentiellement leur attention dessus.
- Les distracteurs (particulièrement sonores) devront donc être limités au maximum pour éviter qu'ils perturbent l'attention des élèves à cause des possibles déplacements attentionnels, particulièrement s'ils ont un bas seuil d'activation.

1.2 Attention divisée

L'attention divisée correspond à la capacité à traiter plusieurs informations en même temps. Néanmoins, plusieurs conceptions ont été énoncées et sont encore débattues pour tenter de comprendre comment nous faisons pour effectuer plusieurs tâches en même temps. Un point de divergence est lié au nombre de ressources attentionnelles dont nous disposons pour réaliser différentes tâches simultanément.

Kahneman (1973) a proposé l'hypothèse d'une ressource unique en attention. Cette ressource peut être allouée au traitement de différentes tâches à réaliser dans la limite de sa capacité. Si la quantité d'attention requise pour l'ensemble des tâches n'excède pas la quantité disponible d'attention dont nous disposons, alors le partage de l'attention sera efficace. Dans le cas contraire, l'allocation des ressources ne sera pas idéale et entraînera des performances amoindries sur certaines tâches, voire sur l'ensemble des tâches. Kahneman a relevé plusieurs facteurs influençant le partage de l'attention. Parmi eux, il y a la nature de la tâche : les tâches faciles ou familières sollicitent moins d'attention que les tâches nouvelles ou difficiles. Il y a aussi les intentions et objectifs des individus qui influencent ce sur quoi ils vont porter leur attention et l'allocation des ressources attentionnelles. Cela renvoie aux traitements descendants (ou *top-down*) ainsi qu'à la notion de motivation intrinsèque de Deci (1971). En effet, ce dernier propose que les individus ne sont pas

nécessairement motivés par une récompense pour effectuer une activité, c'est-à-dire par une motivation extrinsèque. Ils peuvent être motivés par “*le plaisir qu'[ils] éprouve[nt] par la pratique d'[une] activité ou au travers des sentiments de satisfaction qu'il[s] [ont] en retour de cette pratique*” (Deci, 1971). Par ailleurs, cela peut aussi nous renvoyer au besoin d'autonomie selon la théorie de l'autodétermination qui “*suppose que la personne décide volontairement de son action [...] et qu'elle l'assume complètement*” (La Guardia et Ryan, 2000), ce qui va dépendre de son intérêt pour l'expérience (Ryan et Edward, 2020). Ainsi, nous pouvons imaginer que plus un individu présente un intérêt élevé pour une tâche, plus il va investir de ressources attentionnelles dans celle-ci.

Ensuite, Wickens (1984) a énoncé un modèle selon lequel nous n'aurions pas une ressource attentionnelle unique (comme pour Kahneman) mais de multiples ressources attentionnelles, qui se différencient notamment par la modalité sensorielle des informations à traiter. Ces différentes ressources ne seraient pas sollicitées pour l'ensemble des tâches à effectuer simultanément mais pour certaines d'entre elles seulement. En d'autres termes, si deux tâches se ressemblent, elles vont puiser dans les mêmes ressources attentionnelles. Ainsi, en situation de double tâche, si l'une de ces deux tâches nécessitent davantage d'attention que la seconde, alors la réalisation de la première tâche se fera au détriment de la seconde. À l'inverse, si les deux tâches ne puisent pas dans les mêmes ressources attentionnelles, alors la réalisation de l'une n'influencera pas celle de l'autre, même si l'une sollicite beaucoup d'attention. Par exemple, conduire tout en ayant une conversation avec un passager est possible, dans la mesure où des informations sonores n'impliquent pas les mêmes ressources attentionnelles que des informations visuelles. Cependant, conduire tout en regardant son téléphone sera bien plus compliqué car toutes deux impliquent des informations visuelles.

De plus, le contenu des informations influence la division de l'attention. Plus précisément, plus les contenus de deux tâches se ressemblent, plus le traitement de l'une risque d'interférer avec le traitement de la seconde. Par exemple, il y aura moins d'interférences entre deux tâches si la première implique une réponse manuelle à la suite d'un stimulus visuel (i.e. une tâche visuo-manuelle) et que la seconde implique une réponse vocale à la suite d'un stimulus auditif (i.e. une tâche auditive-vocale), que si les types de réponses sont inversées (i.e. une tâche auditive-manuelle et une tâche visuo-vocale). Dans ce dernier cas, il devient alors plus difficile de diviser son attention, ce qui peut se traduire par une augmentation du temps de réaction (Levy et Pashler, 2001). Ce phénomène est appelé diaphonie ou *crosstalk* (Navon et Miller, 1987). Selon cette proposition, un stimulus est traité en termes de représentations (ou de codes) qui passe d'une étape de traitement à une autre jusqu'à la réalisation d'une réponse. Le contenu de ces représentations ou codes dépend du type de tâches. Lorsque “*notre système cognitif traite des codes qui se ressemblent entre différentes tâches, des interactions peuvent survenir entre*

les contenus des codes de ces tâches” (Maquestiaux, 2017). Ces interactions diminueraient alors les performances dans la réalisation simultanée des deux tâches. Dans l'exemple donné plus haut, les codes des deux tâches sont similaires entre la tâche auditive-manuelle et la tâche visuo-vocale car toutes deux activent des représentations verbales et des représentations visuo-spatiales. Alors que pour les tâches auditive-vocale et visuo-manuelle, les codes sont plus distincts : la première tâche implique des représentations verbales et la seconde des représentations visuo-spatiales, mais pas les deux.

Enfin, la quantité d'informations qu'il nous est possible de traiter en même temps est limitée par des filtres (cf. Partie I.A. *L'attention sélective*) ou goulets d'étranglement qui s'observent lors de protocoles de double tâches (e.g. Levy, Pashler et Boer, 2006). Dans ce type de tâches, les sujets doivent réagir à un stimulus lié à une tâche (e.g. déterminer si un son est aigu ou grave) apparaissant lors de la réalisation d'une première tâche (e.g. être dans un simulateur de conduite). Dans ce type de protocoles, on note un ralentissement cognitif d'au moins 350 ms et qui est appelé effet de la Période Réfractaire Psychologique (ou effet PRP). Une explication de cet effet est que lorsque nous réalisons une tâche, nous devons traiter un premier stimulus. Si lors de ce traitement un second stimulus (lié à une autre tâche) apparaît, le traitement de ce second stimulus sera “mis en attente” tant que le traitement du premier stimulus n'est pas terminé. Il semblerait alors que nous ne puissions pas réellement traiter deux stimuli simultanément et donc qu'il nous est impossible de faire deux choses à la fois.

Quoi qu'il en soit, effectuer différentes tâches “en même temps” implique un mécanisme de contrôle de l'allocation de la (ou des) ressource(s) attentionnelle(s), étant donné la quantité limitée de ressources attentionnelles dont nous disposons. Ce mécanisme est assuré par une structure en mémoire de travail : le contrôleur central.

Au vu des constats ci-dessus, il sera important pour notre projet de faire en sorte que :

- La tâche à réaliser n'implique pas d'interférence dans les types de stimuli à traiter : par exemple, les vidéos ne devront pas contenir de contenu textuel ou imagé différent du contenu audio, au risque d'avoir une interférence.
- Les élèves ne réalisent pas plusieurs tâches en même temps puisque le système attentionnel ne le permet pas, d'autant plus que l'utilisation du *Whisperer* risque de demander beaucoup d'attention.
- Pour cela, les élèves qui utilisent le *Whisperer* devront se concentrer essentiellement sur le contenu audio (avec la vidéo qui l'accompagne) et faire des pauses lorsqu'il s'agit d'effectuer une autre tâche (par exemple, restituer les informations à leurs camarades).

2 Mémoire de travail

2.1 Modèle de Baddeley et Hitch

La mémoire de travail est un registre de mémoire dont la capacité est limitée et qui permet le maintien temporaire des informations tout en assurant le traitement de celles-ci (Daneman et Carpenter, 1980). La capacité de ce registre est estimée à sept (plus ou moins deux) éléments ou groupes d'éléments (les chunks). Le traitement permet notamment aux informations d'accéder au registre de mémoire à long terme pour y être stockées.

Différents modèles de la mémoire de travail ont été énoncés mais les plus connus restent ceux de Baddeley et Hitch (1974) puis de Baddeley (1986). Selon ce modèle, la mémoire de travail serait constituée de plusieurs systèmes esclaves et d'une structure centrale appelée contrôleur central, administrateur central ou encore système central exécutif. Baddeley propose trois systèmes esclaves : la boucle phonologique, le calepin visuospatial et le buffer épisodique. Le premier permet de stocker et traiter les informations verbales (ou phonologiques) alors que le deuxième permet le stockage et le traitement des informations visuo-spatiales. Le buffer épisodique, lui, est un système tampon entre la mémoire de travail et la mémoire à long terme et stocke des "*informations multisensorielles sous une forme qui combine des représentations visuelles, auditives et autres*" (Lemaire et Didierjean, 2018). L'administrateur central a pour rôle d'assurer la coordination et les interactions entre les différents sous-systèmes de la mémoire de travail. Il intervient aussi dans l'allocation des ressources attentionnelles ce qui en fait une structure essentielle pour la réalisation de tâches cognitives de haut niveau comme le raisonnement, le calcul ou la compréhension du langage (Lemaire et Didierjean, 2018). En effet, ce type de tâches implique des sous-activités dans lesquelles il faut répartir nos ressources attentionnelles dans chacune d'elles pour pouvoir les réaliser. Kane et Engle (2003) parlait d'attention exécutive pour nommer le contrôleur central de Baddeley et Hitch. Pour eux, il servirait à résister aux interférences en permettant de maintenir les représentations liées au but à atteindre et aux étapes qui y sont liées (Maquestiaux, 2017). Dans tous les cas, bien qu'il existe différents termes pour définir cette structure centrale, les auteurs semblent s'accorder sur l'idée qu'elle est essentielle à l'attention.

2.2 Lien entre mémoire de travail et attention

Un certain nombre de travaux ont montré l'influence de la mémoire de travail sur l'attention. Cela vient conforter l'idée d'une structure en mémoire de travail intervenant dans cette fonction.

Tout d'abord, il y a les différences interindividuelles au niveau de la capacité de la mémoire de travail. Les individus avec une capacité de mémoire de travail élevée présentent une meilleure capacité de contrôle attentionnel (i.e. la capacité à diriger son attention selon les buts à atteindre) que ceux ayant une faible capacité (Kane, Bleckley, Conway et Engle, 2001). De plus, une grande capacité de mémoire de travail amène à un taux d'habituation à des distractions auditives plus rapide, permettant de mieux résister à celles-ci (Sörqvist, Nössl et Halin, 2012). Autrement dit, les individus apprennent à associer plus rapidement des stimuli non pertinents au fait que ceux-ci n'ont aucun intérêt pour atteindre leurs buts. Ainsi, ces "*stimuli ne retiennent plus l'attention et leur pouvoir de distraction comportementale est diminué*" (Sörqvist, Nössl et Halin, 2012).

De la même manière, une grande charge en mémoire de travail (e.g. comparer le chiffre de gauche d'un nombre avec le chiffre de gauche du nombre d'un précédent essai) réduit la distraction à des sons non pertinents (SanMiguel, Corral et Escera, 2008). Par ailleurs, les données issues des neurosciences soutiennent le lien entre une charge cognitive élevée en mémoire de travail et la résistance aux distracteurs (Klemen et al., 2010). Ces auteurs ont montré qu'une charge élevée en mémoire de travail auditive entraîne une diminution de l'activité du cortex occipito-latéral lors d'une tâche avec des éléments visuels distracteurs. Or cette région intervient dans le traitement visuel. Ainsi, le traitement de ces éléments visuels non pertinents est amoindri et les individus sont plus résistants à leur distraction. De plus, les réponses évoquées auditives du tronc cérébral (ou ABR) à des sons non pertinents diminuent à mesure que la charge de la mémoire de travail augmente (Sörqvist, Stenfelt et Rönnberg, 2012). Il s'agit de l'activité neuronale produite lorsqu'un son est entendu. Ce son est transformé en signal nerveux et est transmis par le tronc cérébral vers le cortex auditif. Ainsi, un ABR plus faible traduit le fait que les sons ne sont pas (ou peu) traités et n'ont donc pas (ou peu) d'effets distracteurs. Pris ensemble, ces résultats aboutissent à une conception "*unifiée de l'attention où la capacité d'un mécanisme central (la mémoire de travail) module les traitements précoces/sensoriels*" (Sörqvist, Stenfelt et Rönnberg, 2012).

En résumé, lorsque notre mémoire de travail est sur-sollicitée, nous n'allons traiter que les informations qui nous semblent pertinentes. Il semble alors qu'il faille faire des choix sur les informations sur lesquelles porter son attention. Ces choix s'opèrent par le système central exécutif et s'expliquent par la capacité limitée de la mémoire de travail. Cette proposition constitue l'un des trois principes d'une théorie de l'apprentissage : la théorie cognitive de l'apprentissage multimédia.

En raison de la capacité limitée de la mémoire de travail, le *Whisperer* devra idéalement :

- Fournir les informations pertinentes uniquement, afin que les élèves n'aient pas trop d'informations à traiter et puissent réaliser la/les tâche(s) demandée(s), bien qu'une surcharge en mémoire de travail leur permettrait de résister à des distracteurs.

- De plus, les élèves n'ayant pas tous la même capacité de mémoire de travail, il est important de tenir compte de ces différences inter-individuelles.

- Ainsi, donner trop d'informations pourrait favoriser les élèves ayant une bonne capacité de mémoire de travail par rapport à ceux qui ont une capacité moins élevée.

3 Les théories de l'apprentissage

3.1 La théorie cognitive de l'apprentissage multimédia

La théorie cognitive de l'apprentissage multimédia soutient l'idée que la conception de messages pédagogiques multimédias (i.e. des messages contenant des mots associés à des images) doit être conforme à la façon dont l'esprit humain fonctionne pour amener à un apprentissage significatif (Mayer, 2005). Ainsi, il s'agit d'une approche centrée sur les apprenants et non pas sur la technologie (Sorden, 2012). Cependant, Mayer estime que la conception que les designers ont de ce type de messages est incompatible avec le fonctionnement cognitif humain. En effet, pour eux l'être humain disposerait d'un système de traitement de l'information passif, dont la capacité serait illimitée et qui ne concernerait qu'un seul canal d'informations (Mayer, 2005). C'est à partir de ces éléments que Mayer a décrit les trois principes fondamentaux de cette théorie dans le but d'améliorer la création des messages multimédias.

Le premier principe correspond au fait que nous disposons d'un système de traitement de l'information à deux canaux selon le type de stimuli à traiter. Cependant, deux théories sont à prendre en compte pour déterminer ce qui est traité dans chaque canal. La première est centrée sur la modalité sensorielle des stimuli présentés. Cela renvoie aux sous-systèmes de la mémoire de travail du modèle de Baddeley et Hitch (cf. Partie II. A. *Le modèle de Baddeley et Hitch*). Ainsi, un canal traiterait les informations présentées visuellement (par le calepin visuo-spatial) alors que le second traiterait les informations auditives (par la boucle phonologique). La seconde théorie est celle de Paivio (2006) basée sur le mode de représentation des stimuli. Les stimuli verbaux seraient traités par un canal et ceux non-verbaux par le second canal de traitement. Ainsi, ce principe souligne l'importance du mode de présentation des informations car selon la modalité sensorielle

concernée, elles ne seront pas traitées de la même manière (e.g. le traitement d'un son n'est pas le même que le traitement d'une image). Pour autant, bien qu'une information puisse entrer initialement par un canal, les individus sont capables de convertir sa représentation vers le second canal (Mayer, 2005). Par exemple, un texte écrit sur un écran peut être traité par le canal visuel mais les individus peuvent convertir ces images en sons. Sons qui seront traités par le canal auditif ce qui permettra de prononcer des mots à voix haute.

Le deuxième principe de la théorie de l'apprentissage multimédia indique que les canaux ont chacun une capacité de traitement limitée. Cette capacité correspond aux limites de la capacité de la mémoire de travail. Dès lors, il est possible d'être surchargé d'informations si le nombre d'éléments à traiter dépasse cette limite. Cette situation se réfère au concept de surcharge cognitive, issue de la théorie de la charge cognitive de Sweller (1988). Selon cette théorie, lors d'une tâche d'apprentissage, nos ressources cognitives se distribuent dans trois types de charge. La première est la charge intrinsèque, relative aux ressources allouées pour manipuler les éléments à apprendre. Elle dépend du nombre de concepts à manipuler simultanément, de leurs interactions et de leur complexité. La charge extrinsèque correspond aux éléments stockés en mémoire de travail qui ne sont pas directement utiles à l'apprentissage (e.g. des éléments distracteurs ou des consignes trop complexes). Le troisième type de charge est la charge utile. Elle désigne les ressources restantes permettant d'intégrer des connaissances dans la mémoire à long terme. La théorie de la charge cognitive souligne l'importance d'atténuer les charges intrinsèques et extrinsèques pour dédier plus de ressources cognitives à l'apprentissage en tant que tel. Ainsi, lors de la conception de messages pédagogiques multimédias, il faut être prudent quant aux activités proposées et à leur attractivité. En effet, ce sont des facteurs qui risquent de surcharger la mémoire de travail avant que "*l'apprenant n'acquière le concept ou la compétence à apprendre*" (Sorden, 2012).

Enfin, le troisième principe met en évidence l'importance d'un processus actif dans l'apprentissage. En effet, les individus ne sont pas comme des dictaphones ou des caméras qui encodent passivement les informations qu'ils perçoivent et doivent apprendre. Ils ont besoin de leur donner du sens et donc de les traiter (Mayer, 2005). Ainsi, un apprentissage actif implique l'exécution d'un ensemble coordonné de processus cognitifs. L'apprentissage débute par la sélection d'informations pertinentes, pour ensuite les organiser en une structure cognitive cohérente dans la mémoire de travail, et enfin les intégrer avec des connaissances antérieures. Ce processus permet de construire un modèle mental complet, offrant la possibilité de manipuler n'importe quelle information. Pour la conception multimédia, ce principe suggère l'intérêt d'avoir un matériel présentant une structure cohérente, et également l'importance d'avoir un message qui guide l'apprenant sur la manière de structurer ses connaissances. De plus, Bonwell et Eison (1991) considèrent que l'apprentissage actif dans le contexte scolaire "*implique que les élèves fassent des choses*

et qu'ils réfléchissent à ce qu'ils font". Ce type d'apprentissage sollicite des processus de pensée comme l'analyse et la synthèse des informations. Cela nécessite alors qu'ils soient impliqués dans les activités qu'ils réalisent, ce qui peut passer par des discussions avec d'autres camarades ou une synthèse des informations (Wolfe, 2006). C'est ainsi que cette auteure propose différentes idées liées à la communication entre les individus pour intégrer cette forme d'apprentissage dans un cadre scolaire. Par exemple, il peut s'agir de débats ou de partager ces idées en binôme. Cette idée peut nous renvoyer au concept d'apprentissage collaboratif que nous verrons par la suite.

En somme, la théorie de l'apprentissage multimédia décrit différents principes liés à la cognition humaine afin de créer des messages pédagogiques multimédias adaptés. Ces messages concernent différentes modalités sensorielles. Ainsi, cette théorie est liée à l'idée selon laquelle les individus apprennent mieux lorsque plusieurs médias sont utilisés par rapport à un seul média. Cette idée est au cœur de la théorie de l'apprentissage multisensoriel.

Pour notre projet, il est préférable de :

- Concevoir le *Whisperer* de manière à limiter au maximum les charges cognitives intrinsèque et extrinsèque afin de dédier plus de ressources à l'apprentissage lui-même.
- Pour cela, nous devons tenter de limiter au mieux les éventuels distracteurs et éviter de surcharger la mémoire de travail.
- De même, il faudra minimiser la complexité de la tâche à réaliser, par exemple, en faisant en sorte que les élèves comprennent instinctivement ce qu'ils doivent faire et comment utiliser le *Whisperer*.
- Enfin, nous chercherons à favoriser l'apprentissage actif en rendant les élèves acteurs de leur apprentissage. Pour cela, les sessions de travail comporteront des périodes durant lesquelles les enfants discuteront et réfléchiront ensemble.

3.2 La théorie de l'apprentissage multisensoriel

Dans le cadre de l'apprentissage scolaire, les informations à apprendre peuvent être traitées par différentes modalités sensorielles, généralement la vue et l'audition. Notre expérience dans le monde implique une stimulation multisensorielle constante. La théorie de l'apprentissage multisensoriel suggère qu'une stimulation multisensorielle serait plus propice à l'apprentissage qu'une stimulation par une seule modalité sensorielle. Historiquement, le système perceptif était considéré comme une fonction modulaire dans laquelle les différentes modalités sensorielles fonctionnent indépendamment les unes des autres. A ce jour, les études suggèrent l'idée d'un système perceptif traitant des interactions multisensorielles. Dans cette optique, les informations provenant de différentes modalités sensorielles sont

rassemblées dans le cerveau pour produire une représentation unifiée et cohérente du monde extérieur. Cela peut nous renvoyer au buffer épisodique du modèle de Baddeley et Hitch (1974). Un apprentissage multisensoriel permet de générer plusieurs codages simultanés laissant une trace en mémoire constituée de plusieurs propriétés. L'encodage, le stockage et la récupération d'informations perceptives sont habitués à fonctionner dans un environnement multisensoriel, et le traitement unisensoriel est généralement moins efficace car il recourt à un mode de traitement artificiel n'utilisant pas le système perceptif de façon optimale (Shams & Seitz, 2008). Les informations provenant d'un apprentissage multisensoriel seraient ainsi mieux intégrées. Les études neuroanatomiques, électrophysiologiques et de neuroimagerie montrent que des interactions multisensorielles peuvent se produire tout au long du traitement perceptif (Ghazanfar & Schroeder, 2006 ; Driver & Noesselt, 2008). L'effet McGurk (McGurk & MacDonald, 1976) illustre cette intégration multimodale : le son « ba » a tendance à être perçu comme « da » lorsqu'il est associé à un mouvement visuel des lèvres associé à « ga ». Ce phénomène montre que la modalité visuelle peut influencer notre perception d'une modalité auditive. L'étude de Seitz et al. (2006) met en évidence un apprentissage perceptif plus efficace chez des personnes bénéficiant d'un entraînement à la fois visuel et auditif, en comparaison à un entraînement uniquement visuel. Ces résultats soutiennent l'intérêt de bénéficier d'un entraînement auditif-visuel pour favoriser l'apprentissage. L'entraînement multisensoriel aurait donc pour conséquence une meilleure intégration des informations.

Dans le cas de ce projet :

- Les élèves apprennent mieux en intégrant les informations par plusieurs modalités sensorielles simultanément selon la théorie de l'apprentissage multisensoriel.
- Ainsi, cette théorie soutient l'intérêt d'utiliser un dispositif offrant des informations auditives en complément d'un support visuel.
- Il faut alors que le contenu des informations sonores soit congruent avec celui des autres informations transmises par le reste de l'interface CARDS.

3.3 L'apprentissage collaboratif

L'apprentissage collaboratif est une approche éducative dans laquelle des groupes d'apprenants (dans notre cas des élèves) travaillent collectivement “*pour résoudre un problème, accomplir une tâche ou créer un produit*” (Laal et Laal, 2012). Il s'agit donc d'une approche qui considère que l'apprentissage est social et passe par la discussion (Gerlach, 1994). Néanmoins, deux facettes de la collaboration doivent être distinguées. La première, aussi qualifiée comme la coopération, peut être définie comme “*la poursuite conjointe d'objectifs convenus par les coéquipiers selon une compréhension partagée des contributions et aboutissants*” (Gulati et al. 2012).

Autrement dit, les acteurs se répartissent différentes tâches qui, additionnées, forment la totalité du projet. Dans cette conception, chacun fait quelque chose de différent. Cela se traduirait dans le cadre pédagogique, par des compétences acquises différentes suivant la tâche réalisée. La deuxième facette, aussi appelée coordination, correspond à “*l’alignement ou ajustement délibéré et ordonné des actions de partenaires pour accomplir des objectifs conjointement déterminés*” (Gulati et al. 2012). Ici, tous les acteurs travaillent ensemble de manière synchrone pour atteindre le même but. Cette conception permet aux acteurs de réajuster leurs efforts au fur-et-à-mesure de l’avancement du projet et des feedbacks qu’ils se fournissent. De même, cela permet une combinaison rapide de leurs ressources. Cela est à relier à un des éléments de base pour que l’apprentissage collaboratif soit efficace, à savoir une interdépendance positive des membres clairement perçue. Il s’agit de l’idée que les “*membres sont obligés de compter les uns sur les autres pour atteindre l’objectif*” (Laal et Laal, 2012), ce qui pourrait passer par une mise en commun des compétences et des connaissances de chacun des membres. En outre, nous pouvons relier cette conception de l’apprentissage avec l’un des besoins fondamentaux de l’être humain selon la théorie de l’autodétermination, à savoir le besoin d’être en relation à autrui. Ce besoin renvoie à l’idée que l’être humain a besoin d’avoir un sentiment d’appartenance et de connexion avec les autres (Ryan et Laguardia, 2000) et pourrait peut-être expliquer l’efficacité de cette approche.

Dans le cadre de ce projet :

- Il faudrait que le *Whisperer* favorise l’apprentissage collaboratif, et particulièrement la coordination pour que tous les élèves acquièrent les mêmes compétences.
- Ainsi, tous les enfants du groupe devront utiliser le *Whisperer* à tour de rôle lors de la réalisation d’une tâche.
- L’apprentissage collaboratif pourrait passer par des discussions ou des synthèses des informations entendues seulement par les élèves utilisant le *Whisperer* à ceux qui ne l’utiliseront pas. Cela pourrait favoriser une interdépendance positive par le partage d’informations complémentaires.
- Faire en sorte que le *Whisperer* favorise le travail en groupe permettrait de satisfaire le besoin de relation à autrui ce qui pourrait augmenter la motivation intrinsèque des élèves à réaliser le travail qui leur est demandé.

4 Bilan

En résumé, une interface tangible et augmentée a été conçue par l’équipe du projet e-TAC, dont l’objectif est d’améliorer l’apprentissage collaboratif. Son utilisation offre un grand potentiel pédagogique (Giraudeau et al., 2019), mais

présente un inconvénient majeur : la lecture de vidéos par plusieurs groupes d'élèves n'est pas synchronisée, ce qui entraîne une ambiance sonore très bruyante, non propice à l'apprentissage. Pourtant, l'utilisation de la modalité auditive constitue un apport favorable à l'apprentissage aux vues de certaines théories de l'apprentissage. Ceci est d'autant plus vrai lorsqu'elle accompagne les informations visuelles déjà présentes dans l'interface (Seitz et al., 2006).

Au regard des théories en psychologie cognitive et des théories de l'apprentissage, le *Whisperer* devrait idéalement satisfaire les conditions suivantes :

- Réduire au maximum les distracteurs, afin que l'attention de l'élève écoutant reste focalisée sur la tâche.
- Se limiter aux informations essentielles pour ne pas surcharger cognitivement l'élève. Il serait bénéfique de réduire au mieux les charges cognitives intrinsèque et extrinsèque pour que l'élève puisse allouer une plus grande part de ses ressources attentionnelles à l'apprentissage. De même, il est important que l'élève réalise une seule tâche à la fois car le traitement des stimuli par le système attentionnel ne se fait pas simultanément.
- Favoriser l'apprentissage actif, en faisant en sorte que l'élève soit impliqué dans son apprentissage, notamment en interagissant avec ses camarades.
- Transmettre des informations pertinentes par rapport au(x) but(s) à atteindre. Le contenu de ces informations devra également être congruent avec les autres informations transmises par l'interface CARDS. Celui-ci vient en complément des supports visuels et ne doit pas surcharger l'élève d'informations.

Nous devons donc créer un dispositif qui vérifie ces conditions de manière à favoriser l'apprentissage actif et la collaboration chez les enfants dans le cadre scolaire, tout en diffusant discrètement des informations sonores. La conception du *Whisperer* sera détaillée dans la partie suivante.

5 Conception du *Whisperer*

5.1 Conception centrée sur les utilisateurs

Selon Abras et al. (2004), la conception centrée sur les utilisateurs “*décrit les processus de conception dans lesquels les utilisateurs finaux influencent la manière dont une conception prend forme*”. Ainsi, il s'agit d'une approche centrée sur les besoins et désirs des utilisateurs dont la visée est de garantir que l'utilisation d'un produit soit celle pour laquelle il a été prévu avec le moindre effort d'apprentissage. Cette approche se différencie de la conception centrée sur l'objet car celle-ci ne tient pas compte du fait que les utilisateurs peuvent inventer des fonctions à un produit selon leurs besoins (Krippendorff, 2004). Il est alors nécessaire que les

utilisateurs soient impliqués tout au long du processus de conception. En effet, d'une part, ils informent les concepteurs sur le(s) contexte(s) d'utilisation d'un produit. D'autre part, les évaluations qu'ils fournissent permettent d'améliorer progressivement le produit par des affinements et des spécifications de conception pour qu'il réponde au mieux à leurs besoins et désirs (Chammas, Quaresma et Mont'Alvão, 2015). C'est en ce sens que la conception centrée sur l'utilisateur est itérative car elle passe par des cycles de conception, puis de création de prototypes et finalement de tests et d'évaluations. Ces dernières peuvent se faire par des simulations, des entretiens, des questionnaires ou des tests utilisateurs (Abrás, 2004). Les tests utilisateurs consistent à "*regarder un ou plusieurs participants effectuer des tâches spécifiques avec le produit dans un environnement de test spécifié*" (Lewis, 2006). Bien que cet auteur considère que ce type de méthodes doit être utilisé avec prudence, il semble partager l'idée de Desurvire et al. (1992) selon laquelle les tests utilisateurs sont la meilleure méthode pour recueillir des données sur l'utilisabilité, c'est-à-dire sur la facilité d'utilisation et l'acceptabilité d'un produit (Bevan, Kirakowski et Maissel, 1991). Ainsi, le prototype que nous proposons, dont les caractéristiques et la forme seront présentées dans les parties suivantes, aura besoin d'être testé auprès des élèves et de leurs enseignants et sera amené à être optimisé.

5.2 Forme du *Whisperer*

Comme nous l'avons énoncé plus tôt, des éléments distracteurs ou des consignes trop complexes augmentent la charge extrinsèque lors d'une activité (cf. Partie III. A. *La théorie cognitive de l'apprentissage multimédia*). Un objet dont l'utilisation n'est pas claire à première vue monopolise des ressources cognitives au détriment du contenu de l'activité pédagogique en cours. Par conséquent, une notion importante à aborder pour réduire la charge mentale extrinsèque que pourrait susciter le *Whisperer* est celle d'affordance. Celle-ci a été initialement proposée par Gibson (1979) et peut être définie comme les possibilités d'action d'un individu dans un environnement particulier. Plus précisément, il s'agit de la "*faculté de l'homme, et de l'animal en général, à guider ses comportements en percevant ce que l'environnement lui offre en termes de potentialités d'actions*" (Luyat et al. 2009). Pour Gibson, une affordance existe par rapport aux capacités d'action d'un individu. Par exemple, un individu ne perçoit pas la hauteur d'un obstacle en centimètres mais il perçoit s'il est capable (ou non) de franchir cet obstacle selon ses propres capacités. Cette définition de l'affordance contraste avec celle fournie par Norman (1988). Celui-ci définit l'affordance comme l'aspect de la conception d'un objet qui suggère comment l'objet doit être utilisé. Cette définition renvoie aux propriétés réelles et perçues d'un objet. Ainsi, Norman considère que la perception d'un individu a une importance dans l'existence d'une affordance. Autrement dit,

pour lui l'affordance renvoie “à la fois à la possibilité d’action et à la façon dont cette possibilité d’action est transmise ou rendue visible à l’acteur” (McGrenere et Ho, 2000). Ainsi, il considère que les affordances permettent de savoir comment utiliser un objet simplement en le regardant, si les concepteurs en tirent parti. Gaver (1991) donne l’exemple d’une poignée de porte (Voir *Figure 2*). Si elle est verticale et fine (Exemple A), alors elle permet de tirer la porte car la configuration de la main et du bras rend plus facile le fait de tirer que de pousser. Si elle est horizontale et plate (Exemple B), alors elle permet de pousser car la configuration de la main et du bras rend plus facile le fait de pousser dans cette situation. Cependant, les affordances perceptives ne correspondent pas nécessairement aux affordances réelles. Par exemple, la poignée de porte verticale suggère de tirer (affordance perceptive ou apparente) mais il est possible que la porte soit verrouillée (affordance réelle). Ainsi, ce que nous percevons peut nous suggérer des possibilités d’actions qui n’existent pas, et inversement. C’est pour cela qu’il faut que les “*affordances apparentes d’un artéfact correspondent à son utilisation prévue*” pour qu’il soit simple à utiliser (Gaver, 1991).

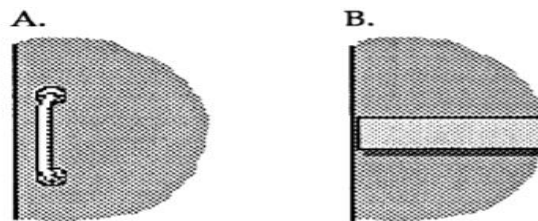


Figure 2 (tirée de l'article de Gaver, 1991) - Différentes poignées de portes suggèrent des affordances pour des actions différentes

Ainsi, la forme du *Whisperer* devrait permettre deux affordances, à savoir celle de préhension ainsi que celle d’écouter (*Annexes 1 et 2*). Pour susciter une action de préhension du dispositif ainsi que d’écoute, nos deux prototypes reconstituent la forme de deux croches. La note de musique est une représentation évocatrice du son qui permettrait d’aider les enfants à comprendre l’utilité de cet objet. Dans un prototype, les deux croches constituent le support sur lequel deux enceintes amovibles se placeraient (*Annexe 1*). Ces enceintes, par leur forme, engloberaient l’ensemble de l’oreille des élèves et offriraient une affordance d’écoute. Dans l’autre prototype, les deux croches constituent dans leur intégralité le dispositif (*Annexe 2*). L’avantage de cette version est que la hampe de la croche, facilement assimilable à un manche, évoque offre une affordance de préhension.

5.3 Caractéristiques du *Whisperer*

Les problèmes liés au bruit sont susceptibles d'être contraignants pour l'apprentissage. Il faut alors pouvoir les réduire. Pour cela, nous pouvons nous intéresser à la réduction de bruit que nous retrouvons dans les casques audios. En effet, l'efficacité de l'utilisation de casques à réduction de bruits a déjà été montrée pour la réalisation de tâches simples et doubles par rapport à l'absence de casques (Molesworth, Burgess et Kwon, 2013). Bien que ces résultats aient été observés dans le cadre de l'aviation, ces auteurs proposent que "*les avantages de cette technologie [peuvent] s'étendre aux situations dans lesquelles l'auditeur peut être distrait ou engagé dans une autre tâche que la tâche cible*". Ainsi, la réduction de bruit pourrait trouver son intérêt dans une interface comme CARDS. Il existe deux types de réduction de bruit : active et passive. La réduction de bruit active consiste à "*annuler activement l'amplitude du son par l'introduction d'une onde sonore similaire*" (Gower et Casalvi, 1994). En effet, les casques ayant ce type de technologie disposent d'un micro qui capte les sons ambiants et qui produit alors des ondes sonores de même amplitude mais qui sont en antiphase. Ces deux types d'ondes s'annulent grâce au phénomène d'interférence destructive ce qui permet de réduire le volume du bruit ambiant perceptible. La réduction de bruit passive, quant à elle, consiste à réduire le bruit ambiant en utilisant des matériaux absorbants ou isolants. En effet, lorsqu'une onde sonore se propage dans l'air et qu'elle rencontre un obstacle, elle est atténuée. Ainsi, plus il y a d'obstacles ou plus ces obstacles sont composés de matériaux isolants, plus l'atténuation est importante. Les casques utilisant ce type de réduction de bruit couvrent souvent l'ensemble de l'oreille pour être le plus étanche possible aux bruits ambiants. Ce type de réduction de bruit est plutôt efficace sur les sons à haute fréquence alors que la réduction de bruit active est efficace sur les basses fréquences. Ainsi, il serait intéressant que le *Whisperer* soit composé de matériaux isolants et qu'il couvre toute l'oreille de l'enfant qui le porte pour favoriser une réduction de bruit passive. Il pourrait aussi contenir un système de réduction active du bruit étant donné que la réduction de bruit passive ne pourrait pas permettre d'isoler, à elle seule, toutes les fréquences sonores liées aux bruits ambiants de la salle de classe.

En outre, le *Whisperer* pourrait présenter un ou plusieurs *signifiers*, terme proposé par Norman (2008) pour désigner "*un indicateur ou un signal dans le monde physique [...] qui indiquent des informations essentielles*" à propos d'un produit. Plus tard, il ajoutera que ce sont des "*signaux perceptibles de ce qui peut être fait*" et qu'ils permettent de rendre visibles des affordances (Norman, 2013).

Ainsi, le *Whisperer* présenterait d'abord un signal lumineux comme une LED qui s'allumerait ou clignoterait seulement lorsqu'il est utilisable. Cela permettrait de rendre cet objet saillant et donc facilement détectable. La saillance d'un objet renvoie à ses caractéristiques visuelles qui le rendent unique par rapport à

d'autres éléments d'une scène visuelle. Cet objet est alors appelé singleton. Dans les tâches de recherche visuelle, la présence d'un singleton amène à un allongement du temps de recherche d'une cible spécifique (Maquestiaux, 2017). De plus, cet auteur parle d'effet *pop-up* pour parler de l'observation faite par Theeuwes (1994), à savoir que l'apparition d'un item distractif avec un léger temps de retard entraîne une augmentation du temps de recherche d'une cible qui se distingue d'items distractifs. Ainsi, "*l'effet pop-up de l'item distractif capture l'attention*", bien que cet item n'ait pas d'intérêt par rapport à la tâche en cours. Nous pouvons donc nous baser sur ce mécanisme pour ajouter au *Whisperer* un élément *pop-up* (ici, une lumière) qui permettrait de capturer l'attention des élèves lorsqu'il leur est possible d'utiliser cet artéfact pour les amener à s'en servir, à la différence des tâches de recherche visuelles classiques où l'élément *pop-up* est à ignorer. En plus du signal lumineux, nous pourrions également ajouter un signal tactile comme une vibration. En effet, il a été démontré que la détection de signaux multisensoriels était plus rapide que celle de signaux concernant une seule modalité sensorielle (Hecht, Reiner et Karni, 2008 ; Huang et al., 2019).

Nous proposons également d'ajouter sur le *Whisperer* un pictogramme dont la forme serait en lien avec le son. Il pourrait s'agir d'une oreille, du pictogramme représentant le volume ou d'un casque (*Figure 3*). Il a déjà été montré que les enfants âgés de 7 à 13 ans ont une bonne compréhension de pictogrammes liés à la consommation de médicaments (Hämeen-Anttila et al., 2004). Ainsi, nous pouvons imaginer qu'un pictogramme lié au son (comme ceux donnés en exemple) peut permettre aux enfants de comprendre que l'artéfact transmet du son. Cela pourrait alors les aider à comprendre comment s'en servir. En effet, utiliser ce type de pictogramme permettrait de rendre visible aux élèves les possibilités d'action de l'artéfact, c'est-à-dire qu'il s'agit d'un objet qui implique d'écouter et donc qu'il doit être mis à l'oreille. Par ailleurs, s'il est efficace, l'utilisation d'un pictogramme pourrait réduire la charge extrinsèque ce qui favoriserait l'apprentissage selon la théorie cognitive de l'apprentissage multimédia).



Figure 3 - Exemples de pictogrammes à mettre sur le *Whisperer*

Conclusion

Cet état de l'art nous a permis d'appréhender des notions importantes pour la conception d'un dispositif tel que le *Whisperer*. Ainsi, nous avons pu dégager les grandes lignes directrices de la conception de cet artéfact et aboutir à des idées de prototypes, qui ne sont que des bases, et sur lesquels nous devons plus largement nous pencher lors de la seconde phase de ce projet. En effet, il sera important de traiter plus en profondeur les interfaces utilisateurs tangibles en nous intéressant à des notions comme celles de métaphores ou d'embodiment (ou cognition incarnée) dont nous aurons besoin pour la suite de la conception du *Whisperer*. Nous devons également préciser les caractéristiques de cet objet, telles que sa taille ou encore sa séquence d'utilisation.

Bibliographie

- Abras, C., Maloney-Krichmar, D., Preece, J. (2004) User-Centered Design. In Bainbridge, W. *Encyclopedia of Human-Computer Interaction*. Thousand Oaks: Sage Publications. (*in press*).
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8., pp. 647-667). New York: Academic Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Clarendon Press.
- Bevan, N., Kirakowski, J., & Maissel, J. (1991). What is Usability? Proceedings of the 4th International Conference on HCI, Stuttgart.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.630.1555&rep=rep1&type=pdf>
- Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom* (ASHE-ERIC Higher Education Rep. No. 1). Washington, DC: School of Education and Human Development, George Washington University.
- Chammas, A., Quaresma, M., & Mont'Alvão, C. (2015). A Closer Look on the User Centred Design. *Procedia Manufacturing*, 3, 5397-5404.
<https://doi.org/10.1016/j.promfg.2015.07.656>
- Corteen, R. S., & Wood, B. (1972). Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology*, 94, 308-313.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in WM and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Deci, E. (1971). The effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18, 105-115.
- Desurvire, H. W., Kondziela, J. M., and Atwood, M. E. (1992). What is gained and lost when using evaluation methods other than empirical testing. In A. Monk, D. Diaper, and M. D. Harrison (Eds.), *Proceedings of HCI '92: People and Computers VII* (pp. 89-102). York, UK: Cambridge University Press.
- Downs, D., & Crum, M. (1978). Processing demands during auditory learning under degraded listening conditions. *Journal of Speech and Hearing Research*, 21(702-714).

Driver, J. and Noesselt, T. (2008) Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Fadeyi, M. O., Alkhaja, K., Sulayem, M. B., & Abu-Hijleh, B. (2014). Evaluation of indoor environmental quality conditions in elementary schools' classrooms in the United Arab Emirates. *Frontiers of Architectural Research*, 3(2), 166-177. <https://doi.org/10.1016/j.foar.2014.03.001>

Fishkin, K. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8(5). <https://doi.org/10.1007/s00779-004-0297-4>

Gaver, W. W. (1991). Technology affordances. In S. P. Robertson, G. M. Olson, & J. S. Olson (Eds.), *Proceedings of the ACM CHI 91 Human Factors in Computing Systems Conference* (pp. 79–84). ACM. <https://doi.org/10.1145/108844.108856>

Gerlach, J. M. (1994). Is this collaboration ? *New Directions for Teaching and Learning*, 1994(59), 5-14. <https://doi.org/10.1002/tl.37219945903>

Ghazanfar, A.A. and Schroeder, C.E. (2006) Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin, New York.

Giraudeau, P., Olry, A., Roo, J. S., Fleck, S., Bertolo, D., Vivian, R., & Hachet, M. (2019). CARDS : A Mixed-Reality System for Collaborative Learning at School. *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*. Published. <https://doi.org/10.1145/3343055.3359721>

Gower, D. W., & Casalvi, J. G. (1994). Speech Intelligibility and Protective Effectiveness of Selected Active Noise Reduction and Conventional Communications Headsets. *Human Factors : The Journal of the Human Factors and Ergonomics Society*, 36(2), 350-367. <https://doi.org/10.1177/001872089403600214>

Gulati, R., Wohlgezogen, F., & Zhelyazkov, P. (2012). The two facets of collaboration: Cooperation and coordination in strategic alliances. *Academy of Management Annals*, 6(1), 531-583.

Hämeen-Anttila, K., Kemppainen, K., Enlund, H., Bush Patricia, J., & Marja, A. (2004). Do pictograms improve children's understanding of medicine leaflet

information ? *Patient Education and Counseling*, 55(3), 371-378.
<https://doi.org/10.1016/j.pec.2003.04.006>

Hecht, D., Reiner, M., & Karni, A. (2008). Multisensory enhancement : gains in choice and in simple response times. *Experimental Brain Research*, 189(2), 133-143. <https://doi.org/10.1007/s00221-008-1410-0>

Huang, G., Steele, C., Zhang, X., & Pitts, B. J. (2019). Multimodal Cue Combinations : A Possible Approach to Designing In-Vehicle Takeover Requests for Semi-autonomous Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 1739-1743.
<https://doi.org/10.1177/1071181319631053>

Jamieson, D. G., Kranjc, G., Yu, K., & Hodgetts, W. E. (2004). Speech Intelligibility of Young School-Aged Children in the Presence of Real-Life Classroom Noise. *Journal of the American Academy of Audiology*, 15(07), 508-517.
<https://doi.org/10.3766/jaaa.15.7.5>

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2), 169-183. <https://doi.org/10.1037/0096-3445.130.2.169>

Kane, M. L., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47-70.

Klemen, J., Büchel, C., Bühler, M., Menz, M. M., & Rose, M. (2010). Auditory Working Memory Load Impairs Visual Ventral Stream Processing : Toward a Unified Model of Attentional Load. *Journal of Cognitive Neuroscience*, 22(3), 437-446. <https://doi.org/10.1162/jocn.2009.21204>

Krippendorff *, K. (2004). Intrinsic motivation and human-centred design. *Theoretical Issues in Ergonomics Science*, 5(1), 43-72.
<https://doi.org/10.1080/1463922031000086717>

Laal, M., & Laal, M. (2012). Collaborative learning : what is it ? *Procedia - Social and Behavioral Sciences*, 31, 491-495.
<https://doi.org/10.1016/j.sbspro.2011.12.092>

- Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*, 880-913.
- La Guardia, J. et Ryan, R. (2000). Buts personnels, besoins psychologiques fondamentaux et bien-être: théorie de l'autodétermination et applications. *Revue québécoise de psychologie*, *21*(2).
- Levy, J., & Pashler, H. (2001). Is dual-task slowing instruction dependent. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 862-869.
- Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving: Is there any stopping the psychological refractory period? *Psychological Science*, *17*, 228-235.
- Lewis, J. (2006). Usability testing. In: G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 1275–1316). Hoboken, NJ: John Wiley and Sons.
- Luyat, M. & Regia-Corte, T. (2009). Les affordances : de James Jerome Gibson aux formalisations récentes du concept. *L'Année psychologique*, *109*, 297-332. <https://doi.org/10.4074/S000350330900205X>
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning* *41*, 31-48.
- McGrenere, J., and Ho, W. Affordances (2000) : Clarifying and evolving a concept. In *Proceedings of Graphics Interface*, Montreal, 2000, pp. 1–8.
- McGuck, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. [doi:10.1038/264746a0](https://doi.org/10.1038/264746a0)
- Molesworth, B. R., Burgess, M., & Kwon, D. (2013). The use of noise cancelling headphones to improve concurrent task performance in a noisy environment. *Applied Acoustics*, *74*(1), 110-115. <https://doi.org/10.1016/j.apacoust.2012.06.015>
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*, 56-60.
- Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 435-448.
- Norman, D.A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.

Norman, D.A. (2008). The way I see IT signifiers, not affordances. *interactions*. 15, 6 (2008), 18–19.

Norman, Donald A. (2013): *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books.

Paivio , A. (2006). *Mind and its evolution: A dual coding approach* . Mahwah, NJ: Lawrence Erlbaum.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective : Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860.
<https://doi.org/10.1016/j.cedpsych.2020.101860>

SanMiguel, I., Corral, M. J., & Escera, C. (2008). When Loading Working Memory Reduces Distraction : Behavioral and Electrophysiological Evidence from an Auditory-Visual Distraction Paradigm. *Journal of Cognitive Neuroscience*, 20(7), 1131-1145. <https://doi.org/10.1162/jocn.2008.20078>

Seitz, A.R. et al. (2006) Sound facilitates visual learning. *Curr. Biol.* 16, 1422–1427.

Shaer, O. Hornecker, E. (2009). Tangible User Interfaces : Past, Present, and Future Directions. *Foundations and Trends® in Human–Computer Interaction*, 3(1-2), 1-137. <https://doi.org/10.1561/11000000026>

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11), 411–417. [doi:10.1016/j.tics.2008.07.006](https://doi.org/10.1016/j.tics.2008.07.006)

Shield, B. M., & Dockrell, J. E. (2003). The Effects of Noise on Children at School : A Review. *Building Acoustics*, 10(2), 97-116.
<https://doi.org/10.1260/135101003768965960>

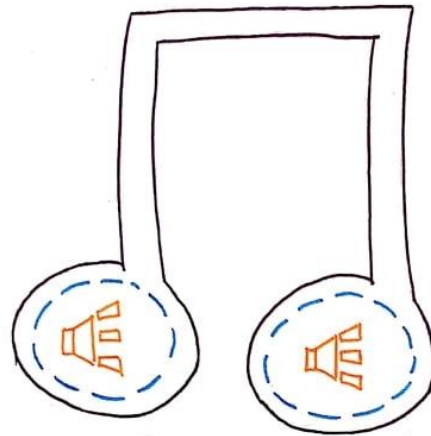
Sorden, S. D. (2012). The cognitive theory of multimedia learning. *Handbook of educational theories 1*, 2012, 1-22.

Sörqvist, P., Stenfelt, S., & Rönnerberg, J. (2012). Working Memory Capacity and Visual–Verbal Cognitive Load Modulate Auditory–Sensory Gating in the Brainstem : Toward a Unified View of Attention. *Journal of Cognitive Neuroscience*, 24(11), 2147-2154. https://doi.org/10.1162/jocn_a_00275

- Sörqvist, P., Nössl, A., & Halin, N. (2012). Working memory capacity modulates habituation rate : Evidence from a cross-modal auditory distraction paradigm. *Psychonomic Bulletin & Review*, *19*(2), 245-250. <https://doi.org/10.3758/s13423-011-0203-9>
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1-28.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285.
- Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 799-806.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*, 242-248.
- Treisman, A. M. (1964). Verbal cues, language and meaning in selective attention. *American Journal of Psychology*, *77*, 206-219.
- Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*, 282-299.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63-102). Orlando, FL: Academic Press.
- Wolfe, K. (2006). Active Learning. *Journal of Teaching in Travel & Tourism*, *6*(1), 77-82. https://doi.org/10.1300/j172v06n01_05
- Zollinger, S. A., & Brumm, H.(2011). The Lombard effect. *Curr Biol* *21*, 16 (Aug 23), R614-615.

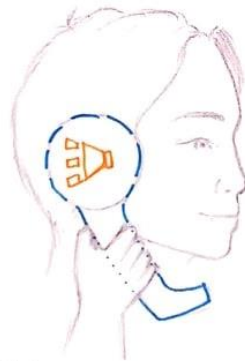
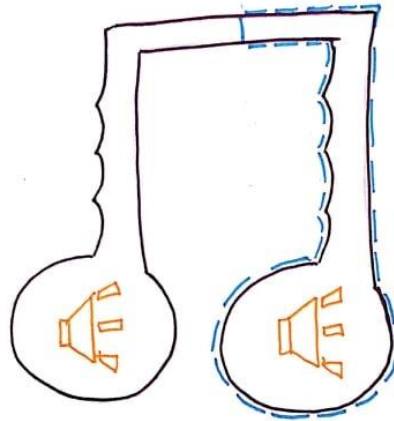
Annexes

Annexe 1 - Premier prototype du *Whisperer*, forme de deux croches avec enceintes amovibles



Scanné avec CamScanner

Annexe 2 - Deuxième prototype Whisperer, forme de deux croches directement préhensibles



CS Scanned with CamScanner