

**Terminal portable de communication et affichage de la reconnaissance vocale. Enjeux et rapports à l'écrit. Étude préliminaire auprès d'adultes déficients auditifs**  
Agnès Piquard-Kipffer, Odile Mella, Jérémy Miranda, Denis Jovet, Luiza Orosanu

► **To cite this version:**

Agnès Piquard-Kipffer, Odile Mella, Jérémy Miranda, Denis Jovet, Luiza Orosanu. Terminal portable de communication et affichage de la reconnaissance vocale. Enjeux et rapports à l'écrit. Étude préliminaire auprès d'adultes déficients auditifs. Ideki 2015 - 3ème colloque international "Didactiques, Métiers de l'Humain, Intelligence collective : construction de savoirs et de dispositifs didactiques", Dec 2015, Colmar, France. pp.1-15. hal-01239910

**HAL Id: hal-01239910**

**<https://hal.inria.fr/hal-01239910>**

Submitted on 15 Nov 2017

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

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

**TERMINAL PORTABLE DE COMMUNICATION  
ET AFFICHAGE DE LA RECONNAISSANCE VOCALE.  
ENJEUX ET RAPPORTS A L'ECRIT.  
Etude préliminaire auprès d'adultes déficients auditifs**

**Agnès Piquard-Kipffer**<sup>1</sup>, Maître de Conférences, Odile Mella<sup>1</sup>, Maître de Conférences, Jérémy Miranda<sup>2</sup>, Ingénieur, Denis Jouvét<sup>2</sup>, Directeur de Recherches, Luiza Orosanu<sup>2</sup>, Doctorante

<sup>1</sup>Université de Lorraine, LORIA, UMR 7503, Vandoeuvre-lès-Nancy, F-54506, France,

<sup>2</sup>Inria, Villers-lès-Nancy, F-54600, France,  
agnes.piquard@loria.fr

**ABSTRACT**

Despite school education and remediation therapy, deaf adults often keep language and communication difficulties with multiple barriers to find employment or access a job corresponding to their abilities. Speech technologies can provide means of helping them by improving their autonomy. This study focuses on an application in French language which has been developed in the collaborative project *Rapsodie* in order to improve communication between a hearing person and a deaf person in working situations. Our principal aim was to investigate different ways of displaying the result of speech recognition transcription which took also into account the reliability of the recognized items. 10 deaf adults have been interviewed. We will present their preferences for displaying the recognition output and examine their links with written language cultural practices. Finally, we will point out their necessary lifelong learning strategies and not only learning throughout their schooling.

**RESUME**

Les aides technologiques, en particulier dans le domaine de la reconnaissance de la parole ambitionnent de favoriser l'inclusion sociale des personnes sourdes. Cette étude examine une application en langue française, développée dans le cadre du projet *Rapsodie*, dans le but d'améliorer la communication entre une personne sourde et un interlocuteur entendant, en situation professionnelle. L'objectif principal était d'étudier différentes possibilités d'affichage des résultats de la transcription du signal de parole prenant en compte la fiabilité des mots et des syllabes reconnus. 10 personnes sourdes ont été interviewées. Les choix d'affichages sont en lien avec leurs modes de communication préférentiels.

## INTRODUCTION

Dans le monde, 70 millions de personnes souffrent d'une déficience auditive. En France, 11% de la population présente ces difficultés (Haeusler, De Laval et Millot, DREES, 2014). Depuis la loi n°2005-102 du 11 février 2005 *pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées*, un droit à compensation est posé comme principe. Il intervient en contexte scolaire et aussi tout au long du parcours de vie des personnes. La perte auditive entraîne des difficultés linguistiques, cognitives et sociales pérennes malgré la prise en charge des enfants sourds de plus en plus précoce (Dumont 2008 ; Expertise collective Inserm, 2007 ; Leybaert, 2005 ; Marschak et Spencer, 2010, 2011). Le manque de moyens de communication partagés fait obstacle à l'intégration et à l'inclusion des personnes sourdes dans un environnement professionnel entendant (Metzger et Barril, 2004). Le taux d'activité de la population en situation de handicap<sup>1</sup> dans lesquelles s'inscrivent des personnes sourdes, a été établi par la Dares comme se portant à 54% en 2013, contre 72% dans la population normale de 15 à 64 ans (Barhourmi et Chabanon, 2015).

Les technologies de l'information et de la communication s'inscrivent dans ce rôle compensatoire des difficultés. Trois types d'aides techniques sont définies par l'OMS : les aides grand public, les produits adaptés, et les aides dédiées. Notre étude fait partie de cette troisième catégorie, s'inscrivant dans le projet collaboratif RAPSODIE (<http://erocca.com/rapsodie>), répondant aux recommandations établies à la suite du rapport de l'ONFRITH (2010) et du rapport Thoumié (2003). Il s'agit, à long terme, de favoriser l'autonomie des personnes sourdes et leur inclusion sociale, leur donnant des moyens de mieux comprendre le discours leur étant adressé avec l'aide d'un système de reconnaissance vocale.

L'objectif général du projet est le développement d'un terminal

---

<sup>1</sup> Cela repose sur la loi de 2005 tenant compte de l'individu et son environnement. Selon la Dares, la population handicapée recouvre à la fois les personnes dont le handicap est reconnu administrativement au sens de la première définition et celles dont le handicap n'est pas reconnu administrativement mais qui déclarent un problème de santé durable accompagné de difficultés dans les tâches quotidiennes.

portable de communication, utilisé en contexte professionnel, dans cette étude, un magasin de bricolage, permettant à une personne sourde ou malentendante, d'interagir avec une personne entendante, cela, sans l'aide d'un interprète. Notre étude recherche le meilleur affichage de la transcription écrite de la reconnaissance vocale sur le terminal, cela en temps réel.

Certaines difficultés sont présentes : les personnes sourdes, pour compenser leurs pertes auditives, doivent pouvoir se reposer sur un niveau fonctionnel de français écrit pour tirer bénéfice de ce terminal de communication. Or, la maîtrise de l'écrit ne va pas de soi pour tous. Le rapport Gillot (1998), a fait état de 80% de personnes sourdes en situation d'illettrisme, sans que la source de ce chiffre n'ait été précisée. Bien que l'ampleur de ces difficultés ait été discutée (Renard, 2001), le niveau de lecture des personnes sourdes ressort dans les études, comme étant inférieur au niveau de lecture des entendants (Expertise collective Inserm, 2007 ; Leybaert, 2005, Marschak et Spencer, 2010, 2011 ; Wang et Williams, 2014), étant précisé que les moyens d'instruction et d'évaluation sont centralement des outils mettant en jeu la phonologie (Allington, 2002 ; Gabriel et Dostal, 2013). Or, la population sourde ne représente pas une communauté homogène. Le moment d'apparition de la perte auditive, son origine et son ampleur, les moyens de communication utilisés sont différents, ainsi que les rapport entretenus avec la langue orale, écrite et gestuelle. Certains adultes, nés sourds, considèrent qu'ils font partie d'une minorité culturelle qui utilise la langue des signes (Fellinger, Holzinger et Pollard, 2012). D'autres adultes sourds, oralistes, considèrent leur perte auditive comme relevant d'un problème de santé. Des obstacles au sein des interactions avec des personnes entendantes sont largement répertoriés. D'autres difficultés sont posées lors du processus de reconnaissance vocale. Des erreurs du système peuvent apparaître, en raison notamment du traitement en temps réel sur un terminal portable. De plus, des bruits peuvent parasiter la parole (musique, bruits de fond) générant une perte d'information et des difficultés de compréhension.

Plusieurs tentatives de développement ont déjà été menées, comme le projet LIPCOM expérimenté dans un contexte scolaire (Coursant-Moreau et Destombes, 1999). Cependant, à notre connaissance, il n'existe pas d'affichage de la reconnaissance vocale d'un interlocuteur entendant, dédié aux adultes sourdes et validé, en terme de taille, couleur et choix des symboles écrits. Notre étude préliminaire s'inscrit dans cette démarche de vérification de la faisabilité du projet auprès de

personnes sourdes, dans le contexte professionnel d'un magasin de bricolage. Dans les sections suivantes, après une brève description du système de reconnaissance, nous détaillerons les modalités proposées pour l'écriture de la transcription vocale. Nous exposerons ensuite les deux phases du protocole expérimental mené auprès de 10 personnes sourdes utilisant des moyens de communication variés. Enfin, nous discuterons des choix d'affichage et des facteurs considérés comme étant les plus aidants.

## 1. Système de reconnaissance, contraintes et possibilités

### 1.1. Système de reconnaissance de la parole

Le système utilisé se décompose en plusieurs phases (cf. Figure 1). Après une analyse acoustique du signal de parole, le moteur de reconnaissance détermine la séquence la plus probable d'unités lexicales prononcées<sup>2</sup>. Ce dernier utilise un modèle statistique de langage hybride qui comprend un lexique de plusieurs milliers de mots et de syllabes. Ainsi, il répond mieux à la contrainte du temps réel et à l'usage d'un terminal portable, utilisant peu d'espace mémoire (Cf. corpus ETAPE<sup>3</sup>, Gravier, Adda, Paulson et al., 2012). Les mots et syllabes sont sélectionnés selon leur fréquence et occurrence dans un corpus de journaux, d'émissions radiophoniques et télévisuelles. Les mots les plus fréquents sont reconnus comme des mots et les mots n'appartenant pas au vocabulaire du système le sont comme des séquences de syllabes (Orosanu & Jouvét, 2014 ; Piquard-Kipffer, Mella, Miranda et al., 2015).

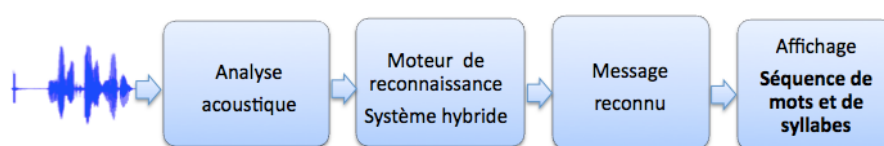


Figure 1. Système de transcription de la parole

### 1.2. Problématique des erreurs de reconnaissance

---

<sup>2</sup> Une unité lexicale est définie ici comme l'entrée d'un lexique (suite de graphèmes correspondant à une suite de sons)

<sup>3</sup> Utilisé pour la transcription d'émissions de radio et de télévision (corpus ETAPE – 82000 mots), plus de 94% des unités lexicales reconnues sont des mots, environ 6% sont des syllabes. Environ 70% des mots reconnus par le système sont corrects ainsi que 60% des syllabes.

Les conditions d'utilisation de ce système de reconnaissance dans un magasin de bricolage sont très différentes de celles ayant servi à l'apprentissage des modèles acoustiques. L'environnement bruyant, la voix du locuteur singulière (articulation, accent, hésitations..), l'intensité vocale, trop forte ou trop faible, peuvent gêner le traitement du signal. Parfois, le système peut ne pas parvenir à reconnaître l'unité lexicale prononcée, même lorsqu'elle fait partie de son vocabulaire. Elle peut également ne pas y figurer. Dans ce cas, elle sera remplacée par une ou plusieurs autres, acoustiquement similaires ; par exemple, le mot *inventeur* reconnu à la place *d'un vendeur*. Ces erreurs du système peuvent gêner l'accès à la compréhension de la phrase affichée sur l'écran du terminal des personnes sourdes.

### 1.3. Possibilités de mise en évidence des erreurs

Il apparaît donc nécessaire de trouver des moyens de faciliter la lecture des phrases reconnues par le système. Une des possibilités est le recours à *la mesure de confiance*. L'objectif est de réduire les ambiguïtés en aidant l'utilisateur à détecter les erreurs commises par le système, cela dans le but de les corriger et de pouvoir accéder à la compréhension. Le système estime le degré de confiance attribuable à chaque mot ou syllabe reconnus. Cette mesure, basée sur des calculs de probabilité *a posteriori* (Jiang, 2005), est une valeur (cf. figure 2). Lorsqu'elle est proche de zéro, l'unité reconnue est considérée comme étant probablement incorrecte (*mesure de confiance faible*). Réciproquement, lorsqu'elle est proche de 1, l'unité reconnue est considérée comme probablement correcte (*forte mesure de confiance*). Toutefois, cette mesure n'est pas totalement fiable, n'étant qu'une estimation calculée par le système lui-même. Des mots mal reconnus peuvent être considérés à tort comme corrects et, réciproquement, des mots bien reconnus, considérés à tort comme incorrects.

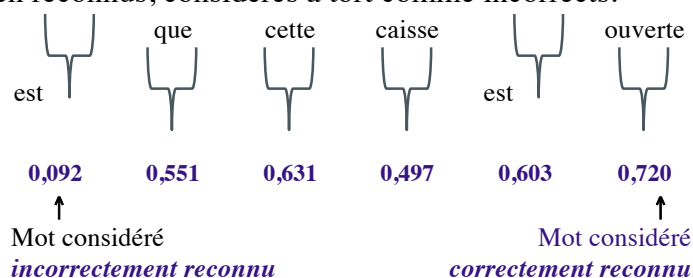


Figure 2. Valeurs de la mesure de confiance comprise entre 0 et 1

Quoi qu'il en soit, cette caractérisation des mots (correctement vs incorrectement reconnus) sera affichée sur le terminal afin d'examiner si cela constitue une aide pour les personnes sourdes.

## 2. Expérimentation 1

L'objectif de cette première phase expérimentale était de trouver l'affichage le plus adapté aux personnes sourdes, indépendamment de la précision du système (cf. Tableau 1). Nous avons émis l'hypothèse selon laquelle la transcription orthographique des phrases ne représenterait pas nécessairement la meilleure écriture pour tous. Tout d'abord, l'écran du terminal portable, après le traitement du signal, affiche une chaîne de mots et de syllabes qui ne peuvent être tous écrits en écriture orthographique. De plus, les niveaux de maîtrise du langage et leurs usages sont inégaux au sein de notre population.

**Tableau 1. Phases expérimentales**

Sélection	Expérimentation 1	Expérimentation 2
Niveau d'audition, de lecture & entretiens	Ecritures d'affichage & entretiens	Avec seuils de confiance & entretiens

C'est pourquoi trois modalités d'écriture ont été étudiées : orthographique, en alphabet phonétique international (API) et en phonétique simplifiée. Dans l'écriture orthographique, les mots reconnus sont écrits selon l'orthographe française et les syllabes en phonétique simplifiée. L'API transcrit les phonèmes pour les mots et les syllabes reconnus. Certains adultes sourds ont bénéficié et/ou continuent à bénéficier, depuis leur enfance, d'entraînements auditifs et langagiers s'appuyant sur ce codage. La troisième modalité d'écriture utilise la phonétique simplifiée. C'est un alphabet conçu pour permettre une lecture aisée de la transcription des sons du langage. Il pourrait être adapté aux personnes sourdes familières de la prononciation de la langue française mais qui ne maîtrisent pas les conventions de l'API. Les phonèmes des mots et syllabes reconnus sont transcrits à l'aide de groupes de lettres de l'alphabet français indiquant plus clairement leur prononciation (par exemple « eu » et « y », respectivement pour /ø/ et /j/ en API).

### 2.1. Méthode

#### 2.1.1. Participants

Les participants sont des adultes sourds bien intégrés socialement, en activité professionnelle ou impliqués dans la vie associative. Le

premier critère d'admissibilité à cette étude incluait un niveau de littéracie fonctionnel, les tâches demandées allant requérir la lecture de la transcription faite, donc l'accès à la compréhension écrite. Le second critère concernait la présence d'une perte auditive d'au moins 45dB (seuils moyens d'acuité)<sup>4</sup> dans la meilleure oreille et l'absence de tout autre difficulté grave qui aurait pu les empêcher de participer aux évaluations. Nous avons interviewé 20 personnes intéressées par ce projet, sur plusieurs semaines, à l'aide d'associations de personnes sourdes ou devenues sourdes, cela sur cinq départements.

La lecture d'un court texte mettant en scène des situations de communication se déroulant dans un magasin de bricolage a été proposée. Chaque participant devait comprendre le rôle d'employé qu'il aurait à jouer, l'interviewer jouant le rôle du client. Cette lecture était ensuite reformulée, pour en vérifier la compréhension. 10 personnes sourdes ont finalement été retenues, n'étant toutefois pas représentatifs de la population sourde. Il s'agit de 4 femmes et de 6 hommes, âgés de 25 à 63 ans, d'une moyenne d'âge 39 ans (cf. Tableau 2). 4 personnes présentent une déficience auditive profonde (de 95-100 dB), 6 personnes une déficience sévère (de 70 à 90 dB, cf. OMS, CIM-10, 1993- WHO-2015). 9 participants utilisent régulièrement des prothèses auditives dans le but d'améliorer leurs informations acoustiques et pour signaler leur particularité. L'âge d'apparition de leurs surdités est variable, les causes étant généralement inconnues. Leurs moyens de communication sont divers. 5 participants, entendants devenus sourds, oralisent. Ils signalent que leur langue maternelle est le français. 2 personnes sourdes de naissance gesticulent et indiquent que la langue des signes française (LSF) est leur langue maternelle. 3 participants, malentendants de naissance devenus sourds ou présentant une surdité diagnostiquée après l'acquisition du langage, combinent des signes issus de la LSF et le respect de l'ordre syntaxique de la langue française (français signé). Ils ne considèrent ni la LSF, ni la langue française comme étant véritablement leur langue maternelle, puisque présentant d'importantes difficultés dans chacune des deux. Nous avons

---

<sup>4</sup> La déficience auditive incapacitante, établie par l'OMS correspondant à une perte d'audition supérieure à 40dB dans la meilleure oreille (pour des sons purs)



relevé leur mode de communication utilisé préférentiellement, celui-ci n'étant pas exclusif.

**Tableau 2. Distribution des 10 participants selon leur mode préférentiel de communication (moyenne-min-max)**

	LSF	Français parlé	Français signé
Age	38 (25-53)	41 (25-64)	38 (28-45) 75
Perte dB	97 (95-100)	77 (70-100)	(70-90)
Sexe M	1	3	1
Sexe F	1	2	2

### 2.1. 2. Tâche et procédure

Les participants devaient lire dix phrases, transcrites par le système de reconnaissance. Nous les avons élaborées selon des critères lexicaux, syntaxiques et sémantiques. Les champs lexicaux étaient ceux du bricolage et de la demande d'information commerciale. D'un point de vue syntaxique, chaque phrase comprenait une ou plusieurs propositions<sup>5</sup>. La cohérence a été respectée. La longueur moyenne était de 11,35 mots (minimum : 5 mots, maximum : 22 mots), afin de ne pas surcharger la mémoire et de faciliter l'accès à la signification. Les phrases étaient en majorité de type interrogatif mais également de type déclaratif, impératif et exclamatif. Elles ont été enregistrées par le même locuteur dans une pièce calme, puis traitées par le système et transcrites. Elles ont ensuite été sélectionnées en fonction de leurs difficultés (présence d'erreurs produites par le système). Enfin, différentes écritures ont été proposées (cf. Tableau 3).

**Tableau 3. Expérimentation 1 – Ecritures d'affichage**

Écritures d'affichage	Résultats de la transcription automatique de la parole en mots et syllabes
Orthographique	je voudrais être li vré combien ça kou te
API	ʒə vudʁɛ ɛtʁ li vʁɛ kɔ̃bjɛ sa ku tə

<sup>5</sup> Constituant de la phrase comprenant un sujet et un groupe verbal.

Phonétique simplifiée	je voudre ètr li vré konbyin sa kou te
-----------------------	--

Les participants ont été vus individuellement. Ils n'ont pas été aidés par des indices sonores, ayant à lire la transcription écrite. Le temps de réponse n'était pas minuté. Nous avons sensibilisé les participants à la présence d'éventuelles erreurs de reconnaissance du système pour plusieurs raisons. Nous ne souhaitons pas que le système de reconnaissance puisse être considéré comme un outil achevé alors qu'il est toujours en évolution. De plus, nous désirions échanger avec les participants des points considérés comme les plus aidants pour faciliter leur compréhension, en présence ou non d'erreurs du système.

## 2.2. Résultats – Expérimentation 1

L'écriture orthographique a été préférée par 9 participants sur 10, spécifiant sont caractère aidant, excepté dans le cas d'erreurs de reconnaissance du système. Ainsi, par exemple, les mots *l'orient* transcrits pour les mots *le rayon* ont gêné les participants, les entraînant vers un champ sémantique inopérant pour asseoir des stratégies de compréhension pertinentes. L'écriture en API n'a été retenue par aucun participant (cf. Tableau 4).

**Tableau 4. Ecritures d'affichage - Préférences des participants (N=10)**

Ecritures d'affichage	Score Total /10	Français		
		LSF (n1=2)	Parlé (n2=5)	Signé (n3=3)
Orthographique	9	2	4	3
API	-	-	-	-
Phonétique simplifiée	1	-	1	-

8 d'entre-eux ont signalé ne pas en être assez familiers. Les 2 autres participants, l'un oraliste et l'autre oralisant et gestuant, l'utilisent en séances d'orthophonie pour coder des mots isolés. Toutefois, selon eux, la lecture de phrases écrites en API mobilise de trop de ressources attentionnelles et cognitives pour permettre un accès précis et rapide à la signification.

L'écriture en phonétique simplifiée a été préférée par un participant, à la fois pour les mots et les syllabes. Etant oraliste, il a indiqué que le passage par la prononciation des mots lui permettait d'accéder au sens

plus facilement, que le système produise des erreurs ou non. Enfin, 7 participants, 5 oralistes et 2 personnes gestuant et oralisant ont suggéré que l'écran propose plusieurs modalités d'affichage. 6 d'entre eux souhaitent d'abord voir afficher l'écriture orthographique, un participant préfère l'ordre inverse.

### 3. Expérimentation 2

L'objectif de cette seconde phase expérimentale était de trouver le meilleur moyen d'afficher des informations complémentaires fournies par le système. La *mesure de confiance* a été utilisée comme indicateur de la correction des unités lexicales. Dans ce but, quatre modalités ont été évaluées (cf. Tableau 5). Les unités lexicales considérées *incorrectement reconnues* par le système ont été présentées en couleur rouge. Les unités lexicales considérées *correctement reconnues* ont été affichées dans la même couleur que les autres unités de la phrase mais en gras.

#### 3.1. Méthode

##### 3.1.1. Tâche et procédure

Deux sessions de tests ont été conduites. En premier, nous avons utilisé une mesure de confiance *oracle* : les unités lexicales considérées comme *incorrectement reconnues* étaient réellement reconnues de manière erronée par le système. Respectivement, les unités lexicales considérées comme *correctement reconnues* étaient également réellement reconnues comme telles. Dans un deuxième temps, nous avons utilisé les mesures de confiance calculées par le système pour classer les unités lexicales reconnues. Nous avons également proposé, pour chacune des modalités d'affichage, des séparations entre les unités lexicales variant de 1 à 3 espaces caractères, afin que le lecteur trouve la meilleure perception possible des syllabes et des mots.

**Tableau 5. Quatre modalités d'affichage pour la phrase prononcée « je voudrais être livré, combien ça coûte ? »**

	Mots/syllabes considérés <i>incorrects</i> affichés en couleur (rouge)	Mots/syllabes considérés <i>corrects</i> affichés en <b>gras</b>
Mots considérés <i>incorrects</i> affichés en mode <b>orthographique*</b>	Je voudrais être <b>li vré</b> <b>qu'on bien ça kou te</b>	<b>je voudrais être li vré</b> <b>qu'on bien ça kou te</b>

Mots considérés <i>incorrects</i> affichés en phonétique simplifiée*	je voudrais être <b>li vré</b> <b>kon byin</b> ça <b>kou te</b>	<b>Je voudrais être li vré</b> <b>kon byin ça kou te</b>
--	--	---

\* Les syllabes sont toujours affichées en pseudo-phonétique .

### 3.2. Résultats – Expérimentation 2

Concernant les mesures de confiance, les choix des participants ont été les mêmes (*oracle* ou mesures du système). La possibilité de mettre en valeur les unités lexicales *correctement reconnues* en bleu a été préférée par l'ensemble des participants. Ils ont signalé que leur attention se portait ainsi sur les mots caractérisés comme corrects (même si certains d'entre eux peuvent être faux), permettant d'asseoir leurs stratégies de compréhension sur des objets ou concepts bien identifiés (cf. Tableau 6). A l'inverse, la présentation en rouge des erreurs - ou supposées telles - de la machine a été clairement rejetée, décrite comme déclenchant une focalisation de l'attention sur des points ne permettant pas d'améliorer un cheminement vers la signification.

**Tableau 6. Préférences des participants pour les modalités d'affichage (N=10)**

Modalité d'affichage *	Mots/syllab les reconnus comme <i>corrects</i> en gras	LSF (n1=3)	França is Parlé (n2=5)	França is Signé (n3=3)
Mots considérés <i>incorrects</i> affichés en mode orthographique	<b>Je voudrais être li vré qu'on bien ça kou te</b>	2	-	-
Mots considérés <i>incorrects</i> affichés en phonétique simplifiée	<b>je voudrais être li vré k on by in ça kou te</b>	-	5	3

\* Séparation entre les unités lexicales de 2 espaces caractères

Les 10 participants ont comparé ce mode aux corrections en vigueur dans leur cursus scolaire passé, notamment lors des tâches de dictée, évoquant leurs sentiments de frustration à ne pas pouvoir avoir recours

à une représentation orthographique précise et rapide sans pour autant pouvoir disposer d'axes constructifs.

Tous les participants ont évoqué l'efficacité du surlignage de certains mots, leur rappelant que le système de reconnaissance pouvait se tromper. En cas d'incertitude du système concernant le codage des mots, 8 personnes sur les 10 interrogées (les 5 oralistes et les 3 participants combinant français oral et gestes) ont préféré l'écriture en phonétique simplifiée, par exemple, *sa me di vs ça me dit*.

L'ensemble des participants a opté pour un affichage séparant les mots de deux espaces caractères. 2 personnes (une pratiquant la LSF et l'autre oraliste) ont suggéré que la présence d'autres espaces entre certains groupes de mots symbolisant les pauses dans le discours du locuteur serait également aidante. L'ensemble des participants a proposé que l'écran du terminal affiche les marques de ponctuation, particulièrement les points d'interrogation, permettant ainsi de distinguer les questions des affirmations. Ils ont néanmoins ajouté que d'autres indices visuels leur permettraient d'accéder à la compréhension, comme l'expression du visage du locuteur (froncement des sourcils) ou une attitude corporelle interrogative (posture d'attente, pointage de certains produits).

#### **4. Discussion**

Les conditions d'expérimentation ont été difficiles pour les participants puisqu'ils ne pouvaient recourir à la lecture labiale et n'avaient pas la possibilité de se baser sur leur audition résiduelle, l'objectif de cette étude étant de se rapprocher de situations de communication dans le bruit pour focaliser leur attention sur les affichages de la reconnaissance vocale. L'ensemble des participants a accueilli favorablement ce projet, précisant qu'ils pourraient utiliser un terminal portable de communication en contexte professionnel, émettant toutefois des réserves par rapport aux contraintes techniques, comme les erreurs du système les déstabilisant. Il apparaît donc nécessaire de souligner aux utilisateurs que le système peut commettre des erreurs, cela même dans l'affichage de mots susceptibles d'être erronés. La personne sourde ne peut donc pas faire entièrement confiance à ces affichages mais devrait plutôt les considérer comme un complément d'information.

Les résultats montrent que la majorité des participants a une préférence pour l'écriture orthographique des phrases transcrites, sauf en cas d'erreur du système. Dans ce cas, ils préfèrent le surlignage de mots considérés comme *correctement reconnus* plutôt que l'inverse. Les vives réactions déclenchées suite au surlignage en couleur rouge des mots supposés *incorrectement reconnus* soulignent le coût attentionnel et émotionnel de la recherche permanente d'indices pour accéder à la signification. Les personnes sourdes se montrent encore marquées par des pratiques pédagogiques passées et révèlent leur sensibilité face à un sentiment fréquent d'incertitude « d'avoir bien compris ». En cas d'erreur éventuelle du système, les participants se scindent en deux groupes. Leurs choix d'affichage sur l'écran du terminal sont liés à leurs modes de communication préférentielle. Nous notons que l'écriture en phonétique simplifiée est pertinente pour les 8 participants utilisant largement l'oral pour communiquer. Ces résultats sont semblables à ceux précédemment observés auprès d'une population entendante (Razik, Mella et Fohr, 2011). Cette transcription implique en effet un accès indirect à la signification des mots en ayant recours à la médiation phonologique, segmentant les mots en syllabes pour les sonoriser dans le but d'accéder à la compréhension. Toute absence de graphème dans la transcription peut compliquer cette tâche et la rendre hors de portée pour un participant peu familier de l'oral et des transcriptions écrites correspondantes. Les 2 participants utilisant préférentiellement la langue des signes ont préféré l'écriture orthographique, même pour les mots considérés *incorrects*. Cependant, pour eux, cet affichage n'est pas aidant en cas d'erreurs. Par exemple, lire *qu'on bien* au lieu de *combien* leur demande des ressources attentionnelles conséquentes avec des risques importants de découragement et d'incompréhension, leurs stratégies de compréhension s'appuyant alors centralement sur le contexte pour compenser ces difficultés. La langue des signes n'ayant pas de modalité écrite, le manque de familiarisation avec le codage sonore des mots écrits limite leur compréhension des phrases affichées. L'aide technologique ne bénéficie donc pas à toutes les personnes sourdes de la même manière.

Pour tous les participants, la perception des mots ou des syllabes est décrite comme étant facilitée par une séparation de deux espaces caractères entre eux. Ce résultat va dans le sens des travaux menés

auprès d'une autre population en difficulté de lecture, celle d'enfants dyslexiques (Zorzi, Barbiero, Facchetti et al., 2012), toutefois, nous n'avons pas examiné la vitesse de lecture dans cette présente étude. Certaines questions posées par les participants comme le traitement de la ponctuation, l'enregistrement sur le terminal de la phrase prononcée et de sa possible réécoute en même temps que la lecture /re-lecture de la transcription écrite restent à examiner.

### **CONCLUSION GENERALE**

Les résultats de cette étude préliminaire valident la faisabilité d'un projet de reconnaissance vocale auprès de personnes sourdes. Il s'agit de prendre en compte les contraintes techniques d'un système embarqué, le traitement de la reconnaissance vocale en temps réel ainsi que les particularités linguistiques des participants. Concernant le choix du meilleur affichage sur un terminal portable, l'écriture orthographique est préférée par la majorité des participants. En cas d'erreur éventuelle du système, tous les participants ont considéré la visualisation d'informations permettant d'estimer la fiabilité du mot reconnu, comme aidante, même si cette estimation n'était pas elle-même indemne d'erreurs. Le surlignage des mots considérés comme corrects est privilégié par tous (par exemple en bleu), le surlignage des mots considérés incorrect est vivement rejeté. En cas d'erreurs éventuelles du système, l'écriture en phonétique simplifiée est jugée très aidante par les participants utilisant largement la langue française orale pour communiquer.

Des études complémentaires seront menées afin d'évaluer le gain de cet outil par rapport à des modes de communication classique. Par ailleurs, il serait intéressant de discuter de ces résultats auprès d'une autre communauté, celle des plus jeunes apprenants.

### **BIBLIOGRAPHIE**

Allington, R. L. (2002). What I've learned about effective reading instruction from a decade of studying exemplary elementary classroom teachers. *Phi Delta Kappan*, 83, 740–747.

Barhoumi M., & Chabanon L. (2015). Emploi et chômage des personnes handicapées. In Synthèse. Stat. Direction de l'animation de la recherche, des études et de la statistique. N°17, novembre 2015. Ministère de l'emploi, de la formation professionnelle et du dialogue social. En ligne :

[http://dares.travail-emploi.gouv.fr/IMG/pdf/synthese\\_stat\\_no\\_17\\_emploi\\_et\\_chomage\\_des\\_personnes\\_handicapees.pdf](http://dares.travail-emploi.gouv.fr/IMG/pdf/synthese_stat_no_17_emploi_et_chomage_des_personnes_handicapees.pdf)

Dumont A. (2008). *Orthophonie et Surdit . Communiquer, comprendre,*

- parler*. Masson, collection orthophonie, 2008.
- Coursant-Moreau A., & Destombes, F. (1999). LIPCOM, prototype d'aide automatique à la réception de la parole par les personnes sourdes. *Glossa*, 68, 36-40.
- Expertise collective Inserm (2007). Chapitre 5. Apprentissage du langage écrit chez les sourds. In *Dyslexie, dysorthographe, dyscalculie. Bilan des données scientifiques*, Editions INSERM, Paris. 137-147. En ligne : <http://www.ipubli.inserm.fr/bitstream/handle/10608/110/?sequence=14>
- Fellinger J., Holzinger D., & Pollard, R. (2012). Mental health of deaf people. *The Lancet*, 379, 1037-1044.
- Gabriel R.E., & Dostal H. M. (2013). « Reading and deafness »: a review. *Read Writ*, 26, 317–320.
- Gillot, D. (1998). « Le droit des sourds : 115 propositions », rapport au Premier ministre, 1998. En ligne : <http://www.ladocumentationfrancaise.fr/rapportspublics/984001595/index.shtml>
- Gravier G., Adda G., Paulson N., Carré M., Giraudel A., & Galibert O. (2012). “The ETAPE corpus for the evaluation of speech-based tv content processing in the French language”, in *Language Resources and Evaluation LREC'12*.
- Haeusler L., De Laval T., & Millot C. (2014). Etude quantitative sur le handicap auditif à partir de l'enquête *Handicap-Santé*, DREES Direction de la recherche, des études, de l'évaluation et des statistiques, *Document de travail, Série Etudes et Recherches*, 131, août 2014.
- Le rapport de L'Observatoire National sur la Formation, la Recherche et l'Innovation sur le Handicap (ONFRIH). *Mars 2010*. 137p. En ligne : [http://travail-emploi.gouv.fr/IMG/pdf/rapport\\_2009.pdf](http://travail-emploi.gouv.fr/IMG/pdf/rapport_2009.pdf)
- Leybaert J., (2005). Reading and Hearing impairment. In M. Snowling, M. Seidenberg, and C.Hulmes (Eds). *Handbook of Reading*, Psychology Press, 379-396.
- Marschak, M & Sprencer, P.E. (Eds) (2010). *The Oxford Handbook of Deaf Studies, Language and Education*. Vol 2. Oxford : Oxford University Press.
- Marschak, M & Sprencer, P.E. (Eds) (2011). *The Oxford Handbook of Deaf Studies, Language and Education*. Vol 1 (2<sup>nd</sup> Edition). Oxford : Oxford University Press.
- Metzger J-L., Barril C. (2004), L'insertion professionnelle des travailleurs aveugles et sourds : les paradoxes du changement technico-organisationnel, *Revue française des affaires sociales*, 58, 3, 63-86.
- Jiang H. (2005), Confidence measures for speech recognition: A survey. *Speech Communication*, 45, 4, 455–470.
- Orosanu L. and Jouvet D. (2014). “Hybrid language model for speech transcription”, in *INTERSPEECH 2014, Singapour*, 2014.



- Piquard-Kipffer A, Mella O., Miranda J., Jouvet D. & Orozanu L. (2015). Qualitative investigation of the display of speech recognition results for communication with hearing impaired people. *6th Workshop on Speech and Language Processing for Assistive Technologies. Dresde 11 september 2015. Interspeech 2015.*
- Razik J., Mella O., Fohr D., & Haton J-P. (2011). Frame-Synchronous and Local Confidence Measures for Automatic Speech recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 25.
- Renard M., (2001). Surdit , accessibilit  et illettrisme. In *Les cahiers de l'actif*, DOSSIER Sourds et malentendants : entre int gration et droit   la diff rence... un conflit d'objectifs ? 298-301, 69-80.
- Thoumi . Ph (2003). Recherche technologique et diffusion de l'innovation au service du handicap. *Minist re d l gu  recherche et nouvelles technologies. Septembre 2003*. 37p. En ligne : [http://www.cofemer.fr/rubrique.php?id\\_rubrique=128](http://www.cofemer.fr/rubrique.php?id_rubrique=128)
- Wang, Y., & Williams, C. (2014). Are we hammering square pegs into round holes? An investigation of the meta-analyses of reading research with students who are d/Deaf or hard of hearing and students who are hearing. *American Annals of the Deaf*, 159 (4), 323–345.
- Zorzi M., Barbiero C., Facoetti A., Lonciari I., Carrozzi M., Montico M., Bravar L., George F., Pech-Georgel C., & Ziegler J. (2012). Extra-large letter spacing improves reading in dyslexia. *PNAS*, 108, 28, 11455-11459.