

DIANDRA FABRE  
INGÉNIEUR, DOCTORANTE  
UNIV. GRENOBLE ALPES, GIPSA-LAB - CNRS, GIPSA-LAB

THOMAS HUEBER  
DR., CHARGÉ DE RECHERCHE AU CNRS  
UNIV. GRENOBLE ALPES, GIPSA-LAB - CNRS, GIPSA-LAB

MÉLANIE CANAULT  
DR., MAÎTRE DE CONFÉRENCES  
UNIV. LYON 1, INSTITUT DES SCIENCES ET TECHNIQUES  
DE LA RÉADAPTATION, LYON, FRANCE  
LABORATOIRE DYNAMIQUE DU LANGAGE, UMR 5596, CNRS  
ET UNIVERSITÉ LYON 2

NATHALIE BEDOIN  
DR., MAÎTRE DE CONFÉRENCES  
LABORATOIRE DYNAMIQUE DU LANGAGE, UMR 5596  
CNRS ET UNIVERSITÉ LYON 2

AUDREY ACHER  
DR., ORTHOPHONISTE  
CHU GRENOBLE ALPES - UNITÉ NEURO-VASCULAIRE,  
PÔLE PSYCHIATRIE-NEUROLOGIE-RÉÉDUCATION

CAMILLE BACH  
ORTHOPHONISTE

LORÈNE LAMBOURION  
ORTHOPHONISTE

PIERRE BADIN  
DR., HDR, DIRECTEUR DE RECHERCHE AU CNRS  
UNIV. GRENOBLE ALPES, GIPSA-LAB  
CNRS, GIPSA-LAB

ADRESSE DE CORRESPONDANCE: PIERRE BADIN  
GIPSA-LAB, F-38000 GRENOBLE, FRANCE  
[pierre.badin@gipsa-lab.grenoble-inp.fr](mailto:pierre.badin@gipsa-lab.grenoble-inp.fr)



# Apport de l'échographie linguale à la rééducation orthophonique

## Contribution of lingual echography to speech therapy



### Résumé

*Les troubles liés à un mauvais positionnement des articulateurs de la parole peuvent avoir des origines diverses, du trouble articulaire isolé à des troubles beaucoup plus lourds liés à une chirurgie intra-orale par exemple. La littérature souligne l'importance de la relation entre la perception et la production de parole, et en particulier celle du retour auditif et proprioceptif, pour l'apprentissage et le maintien du langage. Certaines théories postulent l'existence de modèles ou représentations internes de la parole qui codent les relations entre les mouvements articulaires, leur proprioception, et les résultats acoustiques. Nous faisons donc l'hypothèse qu'une partie des troubles de l'articulation pourrait être attribuée à une dégradation de l'efficacité des canaux sensoriels, et notamment auditifs et proprioceptifs, ou à un défaut des représentations internes du locuteur.*

*Puisque des études ont montré que les sujets sont dotés d'une certaine conscience articulaire et de capacités de lecture linguale, utiliser la vision des articulateurs comme complément aux retours sensoriels pourrait constituer une piste intéressante pour la rééducation des troubles de l'articulation.*

*Dans cet article, nous proposons une revue de la littérature sur l'utilisation de la visualisation des articulateurs en orthophonie. Nous décrivons ensuite une étude pilote portant sur l'évaluation de l'efficacité d'une intervention orthophonique s'appuyant sur l'illustration visuelle échographique des mouvements linguaux pour la rééducation du trouble phonologique de substitution de [tʁ] par [kʁ] chez des enfants âgés de 5 à 7 ans et demi, et montrons que cette approche semble apporter de réels bénéfices à la prise en charge des troubles phonologiques. Ce dispositif présente en outre l'avantage d'être très facile à utiliser dans la pratique d'un cabinet d'orthophonie. Enfin, nous proposons des pistes plus ambitieuses d'évaluation du retour articulaire visuel pour l'orthophonie.*



Orthophonie – Echographie linguale – Proprioception – Trouble de l'articulation – Technologies assistives – Logiciel de rééducation



## Abstract

*Speech disorders related to inaccurate positioning of speech articulators can originate from a persistent speech disorder to much larger disorders due to intraoral surgery for example. The literature emphasises the importance of the link between speech perception and production, and in particular of auditory and proprioceptive feedback, for language acquisition and maintenance. Some theories postulate the existence of speech internal models or representations coding the relations between the articulatory movements, their proprioception, and their acoustic outcome. Therefore, we hypothesise that part of articulatory disorders may be ascribed to some degradation of the efficiency of the sensory channels, in particular the auditory and proprioceptive ones, or to an inaccurate internal representation of the speaker.*

*Since many studies have revealed that subjects are somehow able of articulatory awareness and of tongue reading, using the vision of articulators can be expected to enhance sensory feedback channels and constitute a promising perspective for the remediation of articulatory disorders.*

*In this article, we reviewed the state-of-the-art of the use of articulators' visualisation in speech therapy. Next, we described a pilot study reporting on the evaluation of the efficiency of speech therapy intervention based on the ultrasound visual illustration of tongue movements to remedy a phonological disorder leading to the substitution of [tʃ] with [kʃ] in 5 to 7 years and a half children. The preliminary results provided evidence for actual benefits offered by this approach to the care of phonological trouble. Moreover, this approach appeared very easy to use practically at a speech therapist's office. Finally, we offered more ambitious perspectives for the assessment of visual articulatory feedback in speech therapy.*



Speech therapy – Tongue ultrasound imaging – Proprioception – Articulatory disorder – Phonological disorder – Assistive technologies – Rehabilitation software of specific language impairments – New information technology

## Introduction

La perception des autres et de soi-même est indispensable pour l'apprentissage du langage, comme le notent Perkell (2012) ou Turgeon *et al.* (2015) pour l'audition, ou encore Mills (1987) pour la vision. Ce dernier rapporte les difficultés des enfants non-voyants à apprendre les contrastes entre /m/ et /n/, faciles d'un point de vue visuel, mais difficiles d'un point de vue acoustique. Turgeon *et al.* (2015) montrent plus particulièrement l'importance du retour auditif pour les capacités des locuteurs à compenser la perturbation induite par l'introduction d'un tube entre les lèvres pour la production de la voyelle /u/. Cowie *et al.* (1982) mettent en évidence une tendance des malentendants post-linguaux à une dégradation de la parole qu'ils produisent. Matthies *et al.* (1996) démontrent également l'importance du retour auditif pour le maintien, y compris dans le court terme, des contrastes phonémiques entre /s/ et /ʃ/ pour un sujet malentendant équipé d'un implant cochléaire muni d'une commande marche / arrêt qui permet de l'arrêter momentanément.

Ces liens entre perception et production sont formalisés dans différentes études qui proposent des modèles internes de représentation de la parole. Perkell *et al.* (2000) développent une théorie de la composante segmentale du contrôle moteur. Cette théorie considère que la programmation des mouvements articulatoires pour atteindre des cibles auditives correspondant à des phonèmes ou des segments est basée sur un modèle interne, aussi appelé *copie d'efférence*, qui prédit les conséquences acoustiques à partir des configurations articulatoires. Le retour auditif permet en premier lieu à l'enfant d'acquérir ce modèle interne lors de l'apprentissage du langage, et ensuite de le maintenir tout en l'adaptant aux changements de morphologie liés à la croissance. La copie d'efférence permet d'assurer la correction des commandes motrices à partir des différences entre les cibles acoustiques et les cibles prédites par le modèle en l'absence de retour effectif (Perkell *et al.*, 1997). Les configurations articulatoires sont liées aux informations proprioceptives telles que la longueur des muscles ou les sensations de contact entre organes articulateurs qui font également partie du retour orosensoriel (*Ibid.*).

Comme détaillé dans Badin *et al.* (2010), de nombreuses études ont mis en évidence et quantifié l'apport de la vision des articula-

teurs visibles (lèvres, mâchoire, visage, pointe de la langue, dents) à la perception de la parole (e.g., Erber, 1975; Sumby & Pollack, 1954; Benoît & Le Goff, 1998). Cependant, même si elle est très utile, la seule vision des lèvres et du visage ne fournit qu'une information phonétique très incomplète, en particulier sur les mouvements de la langue ou du voile du palais. Par ailleurs, Montgomery (1981) a montré que l'humain possède une certaine conscience articulatoire, c'est-à-dire une certaine connaissance plus ou moins explicite de la place de ses articulateurs. Cela a incité Badin *et al.* (2010) à tester l'hypothèse selon laquelle les humains seraient capables d'utiliser la vision de la langue pour la reconnaissance des phonèmes, comme ils le font en *lecture labiale*. Cette étude montre que les sujets sont en effet dotés d'une certaine capacité de *lecture linguale* (par analogie avec la *lecture labiale*), c'est-à-dire qu'ils sont capables d'exploiter une représentation visuelle des mouvements de la langue pour améliorer leur perception (notamment des consonnes) lorsque le signal audio est fortement dégradé ou absent. Dans Badin *et al.* (2007), cette représentation visuelle est fournie par un *clone orofacial virtuel* ou *tête parlante articulatoire*, avatar tridimensionnel d'un véritable locuteur (voir ci-après).

Fort de ce cadre théorique sur les liens entre production et perception de la parole, ainsi que sur ces capacités *naïves* à exploiter une information visuelle sur un articulateur dont nous n'avons a priori qu'une représentation très partielle (basée quasi exclusivement sur notre proprioception), plusieurs études se sont intéressées à l'impact d'une représentation visuelle articulatoire pour la rééducation des troubles de l'articulation. On distinguera dans cet article deux paradigmes. Le premier, dit d'*illustration linguale*, vise à fournir au patient une visualisation intuitive d'un mouvement lingual cible, ce dernier pouvant par exemple être enregistré sur un locuteur de référence. Le second, dit de *retour lingual*, vise en revanche à fournir au patient une représentation visuelle de ses *propres* mouvements (et non ceux d'un locuteur tiers), dans son propre espace articulatoire avec sa propre morphologie. L'objectif est ici de l'aider à mieux comprendre et corriger son geste, notamment en le mettant en relation avec ses retours proprioceptif et auditif, et avec les instructions du praticien. Différentes techniques peuvent être mises en œuvre

pour capturer le mouvement de la langue, comme l'électropalato-graphie (voir ci-après). Dans cet article, nous nous intéresserons à l'échographie (ou imagerie ultrasonore), qui semble aujourd'hui être une technique bien adaptée à la rééducation orthophonique.

Dans le cas du retour articuloire par échographie, Ballard *et al.* (2012) discutent deux types de retour (*feedback*) pour l'apprentissage ou la rééducation d'habiletés motrices en parole. Le retour par la *connaissance sur l'exécution (Knowledge of Performance (KP) feedback)*, encore appelé *retour cinématique*, fournit des informations détaillées *externes* sur la façon dont le mouvement a été réalisé, et sur l'écart d'exécution par rapport au mouvement cible en termes d'amplitude ou de qualité. Ces informations permettent ainsi de guider un apprentissage ou une réhabilitation actifs des habiletés motrices. En revanche, le retour par la *connaissance sur le résultat (Knowledge of Results (KR) feedback)* indique seulement si l'objectif de la tâche a été atteint ou non (par exemple un phonème spécifique). Notons que le retour de type KR est inhérent au retour de type KP.

Plus récemment, Blyth *et al.* (2016) rappellent l'hypothèse de Ballard *et al.* (2012) et Newell *et al.* (1990) qui suggèrent que le retour KP est bénéfique quand l'apprenant n'a pas une solide représentation interne du mouvement nécessaire pour atteindre le résultat désiré, comme un but acoustique en parole. Blyth *et al.* (*Ibid.*) font l'hypothèse que c'est le cas des patients ayant subi une glossectomie partielle, qui ont des difficultés à retrouver une parole claire, leur représentation interne ne correspondant plus au nouvel état de leur langue. Ils décrivent une étude sur deux patients visant à évaluer l'utilisation de l'imagerie échographique de la langue comme outil de retour de type KP, et ils établissent un bénéfice thérapeutique du retour visuel par échographie pour la réhabilitation de la parole suite à une glossectomie partielle.

Au vu de ces études, nous faisons l'hypothèse qu'une partie des troubles de l'articulation pourrait être attribuée à une dégradation de l'efficacité des canaux sensoriels, en particulier auditifs et proprioceptifs, nécessaires à l'acquisition ou au maintien des capacités de production de la parole, ou à un défaut des représentations internes du locuteur. Ces différentes déficiences peuvent être liées à une surdit , à un handicap moteur, à une chirurgie des organes oro-

faciaux, ou encore à des problèmes de développement. Utiliser la vision des articulateurs comme un complément de retour sensoriel semble donc une piste intéressante pour la rééducation des troubles de l'articulation.

Dans la suite de cet article, nous présentons l'état de l'art de l'utilisation de la visualisation des articulateurs dans le domaine de la rééducation des troubles de l'articulation, nous décrivons une expérience pilote de rééducation d'un trouble phonologique basée sur l'illustration, et nous proposons des pistes pour l'évaluation du retour articulo-visuel.

## I – Etat de l'art de la visualisation des articulateurs en orthophonie

Nous dressons dans cette section un état de l'art des techniques existantes pour chacun de ces deux paradigmes de représentation visuelle des articulateurs de la parole, l'illustration et le retour.

### A – Illustration pour l'orthophonie

Au détour de collaborations multiples entre ingénieurs, chercheurs et orthophonistes, de nombreux logiciels pour la rééducation des troubles de la parole ont émergé. Parmi les premiers logiciels à proposer une illustration complète des mouvements des articulateurs de la parole, Diadolab permet des animations de contours d'articulateurs d'enfant simplifiés et inspirés de données articulatoires (Menin-Sicard & Sicard, 2012). Dans le même domaine, Canault propose un outil interactif en ligne intitulé *Le conduit vocal en action*<sup>1</sup> modélisant, à partir des données cinéradiographiques de Bothorel *et al.* (1986), les mouvements de l'ensemble des articulateurs impliqués dans la parole (voir une illustration à la Figure 1 gauche). Hueber (2013) présente le logiciel UltraSpeech-player. Ce logiciel<sup>2</sup> s'appuie sur une grande base de films échographiques à haute vitesse de la langue, enregis-

<sup>1</sup> Cette animation peut être testée sur la page <http://anatomie3d.univ-lyon1.fr/webapp/website/website.html?id=3346735pageId=223201>

<sup>2</sup> Ce logiciel est téléchargeable sur la page [www.ultraspeech.com](http://www.ultraspeech.com)



trés de façon synchrone avec le signal de parole et une vidéo également haute-vitesse des lèvres, à l'aide du logiciel UltraSpeech (Hueber *et al.*, 2008). Cette base contient les différentes voyelles et consonnes du français, ainsi que des logatomes de type VCV ou CVC, des clusters de consonnes, des mots et phrases simples, des déglutitions, prononcés quelques locuteurs (dont une orthophoniste). UltraSpeech-player embarque un mécanisme de traitement en temps-réel des flux audio et vidéo, permettant à l'utilisateur de contrôler la vitesse du geste articulatoire présenté, et celle du signal sonore associé. L'objectif est de ralentir le geste articulatoire pour mieux l'observer. L'utilisation de données articulatoires acquises sur un véritable locuteur (et non construite par concaténation de positions statiques), permet de restituer une dynamique réaliste des mouvements, notamment au niveau des patrons de coarticulation. Nous avons pu mettre en application ce logiciel, déjà utilisé par plusieurs orthophonistes, dans le cadre d'une étude qui sera détaillée dans la suite de cet article (voir illustration à la Figure 1 droite).

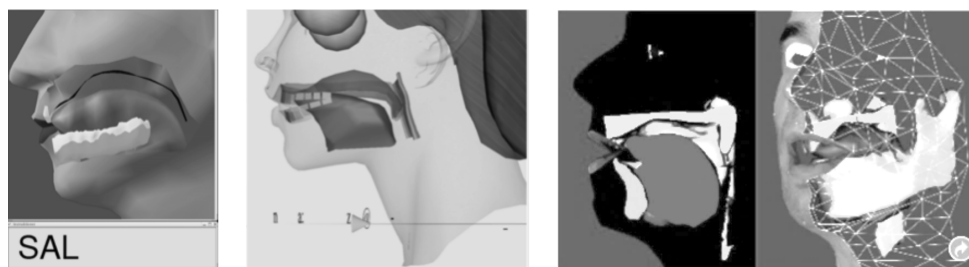


**Figure 1.** Exemples d’affichage avec interface de commande et articulations. *Le conduit vocal en action* (gauche); *UltraSpeech-player* (droite).

Une autre façon de faire visualiser des mouvements articulatoires cibles au patient consiste à utiliser une tête parlante articulatoire, dont la particularité est de rendre possible la visualisation, sous n’importe quel angle, des articulateurs invisibles comme la langue et le voile du palais. Ce type de rendu est classiquement construit à l’aide de techniques de synthèse d’image, s’appuyant sur des modèles articulatoires sous-jacents issus de différentes sources de données articulatoires, acquises sur un locuteur de référence par IRM pour les configurations statiques tridimensionnelles, et par articu-



lographie électromagnétique (voir ci-après) pour la dynamique. Parmi les systèmes existants, on peut citer Massaro & Light (2004), Engwall (2008), Fagel & Madany (2008), ou Badin *et al.* (2010), illustrés à la Figure 2.



**Figure 2.** Exemples de présentation de tête parlante articulatoire. De gauche à droite: Engwall (2008), Fagel & Madany (2008) et Badin *et al.* (2010).

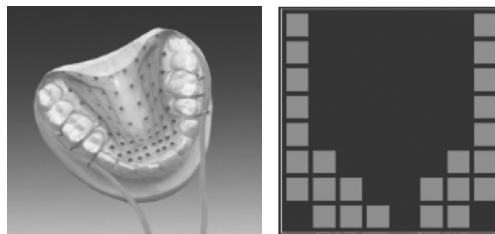
Chen *et al.* (2016) proposent une revue de la littérature sur l'apport de ces techniques d'illustration (une dizaine d'études mettant en jeu une tête parlante y est présentée). Divers modèles de têtes parlantes articulatoires, basés sur des schémas simplifiés ou des modèles plus complexes, sont proposés pour des applications à des populations variées: des personnes malentendantes dotées d'un implant cochléaire, des enfants présentant un sigmatisme, ou encore des personnes aphasiques. L'illustration proposée dans chaque cas semble accélérer l'apprentissage et être facilement assimilée par le patient, qui parvient à associer ce qu'il voit à l'écran, dans l'espace articulatoire d'un autre locuteur, avec sa propre articulation. Eriksson *et al.* (2005) recueillent des informations et des conseils auprès d'orthophonistes et de neuf patients âgés de 8 à 15 ans, pour proposer un éventail de suggestions à mettre en application dans ce domaine. Les personnes interrogées s'accordent généralement à dire que si l'orthophoniste est irremplaçable, il faut cependant pouvoir disposer d'un outil simple et motivant, utilisable en dehors des séances encadrées, et facilement modulable en fonction des besoins des orthophonistes et des troubles de l'enfant.

## **B – Retour articulatoire visuel pour l'orthophonie**

A la différence de l'illustration, décrite dans le paragraphe précédent, le retour articulatoire visuel permet au patient de bénéficier d'un affichage, souvent en temps-réel, de sa propre articulation.

## 1 – Electropalatographie

Parmi les technologies existantes, deux émergent de par leur usage prépondérant en rééducation orthophonique : l'électropalatographie (EPG) et l'échographie ultrasonore. L'EPG utilise un palais artificiel sur lequel est réparti un ensemble d'électrodes (voir Figure 3, gauche). Un dispositif recueille les signaux de ces électrodes et permet de visualiser, en temps-réel, les contacts entre la langue et le palais (voir Figure 3, droite), souvent en affichant côte à côte sur le même écran les contacts du patient et ceux de l'orthophoniste. Cette technologie est largement utilisée au Royaume-Uni, comme par Cleland, McCron & Scobbie (2013).



**Figure 3.** Exemple de palais artificiel pour l'EPG (gauche) et de patient (droite) avec la visualisation de ses contacts palato-linguaux.

## 2 – Articulographie électromagnétique

Malgré sa précision sur les zones de contact langue/palais, l'EPG ne fournit aucune information sur le mouvement de la langue au cours de l'articulation. Pour mémoire, notons que de manière complémentaire, l'articulographie électromagnétique (Electromagnetic Articulography, EMA) fournit en temps réel la position précise de quelques points sélectionnés sur les articulateurs et pourrait ainsi être utilisée pour le retour visuel. Les coordonnées de petites bobines électromagnétiques réceptrices placées en ces points sont déterminées grâce à la mesure de courants induits par des bobines électromagnétiques émettrices fixes. Cette technique est par contre souvent jugée trop invasive pour une utilisation concrète en orthophonie (en outre, les bobines peuvent s'avérer gênantes pour l'articulation de certains phonèmes, en particulier lorsque l'une d'entre elles est placée sur la pointe de la langue). (Katz & McNeil, 2010) proposent cependant une preuve de concept pour des cas d'apraxie de plusieurs patients. Cette étude nécessite un approfondissement quant à son efficacité réelle en comparaison avec des séances traditionnelles.

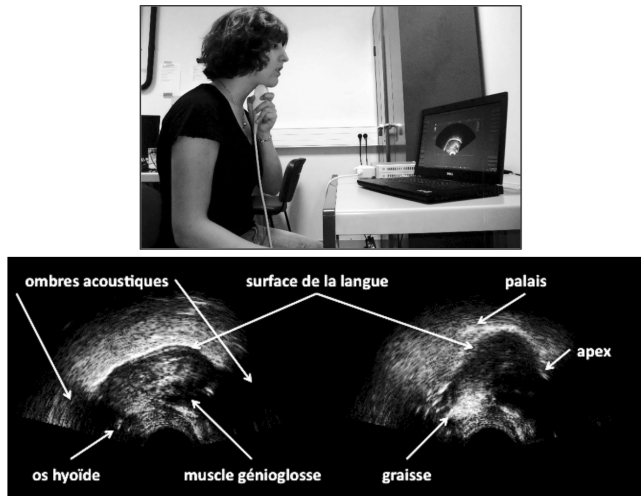
### 3 – Echographie ultrasonore de la langue

L'échographie (ou imagerie ultrasonore) est une technique très intéressante pour observer les mouvements de la langue pendant la production de parole. En effet, cette technique est inoffensive et peu invasive pour le locuteur (Epstein, 2005). Une sonde échographique (médicale) est placée sous la mâchoire du locuteur. Des ondes ultrasonores se propagent dans les tissus de la cavité buccale et sont généralement réfléchies lorsqu'elles atteignent la surface supérieure de la langue. L'analyse de ces réflexions (ou échos) permet de générer une image de la langue. En fonction de la position de la sonde, on obtient une image dans le plan médio-sagittal (plan le plus utilisé, voir Figure 4) ou coronal. Cette technique d'imagerie présente de bonnes résolutions temporelle (de l'ordre de 80 images par seconde dans une configuration standard) et spatiale (résolution inférieure à 1 mm), et ne nécessite pas d'équipement volumineux. Parmi les travaux pionniers sur l'utilisation de l'échographie pour l'étude de la production de la parole, on peut citer Stone & Shawker (1986) ou Stone & Davis (1995). Plus récemment, Hueber & Denby (2009) ont décrit en détail les différents aspects de l'étude du conduit vocal par échographie (concepts physiques sous-jacents, fonctionnement d'un système d'échographie, protocole expérimental pour l'acquisition de données articulatoires, interprétation et traitement des images ultrasonores).

L'échographie semble aujourd'hui être une technique privilégiée dans le cadre du retour visuel pour la rééducation de certains troubles de l'articulation. Les travaux existants sont variés, sur la forme comme sur le fond. Dès 1985, Shawker & Sonies proposent d'utiliser des images échographiques de la langue pour la rééducation du /r/ anglais, avec une patiente âgée de 9 ans. Un enregistrement de l'articulation correcte est joué en boucle conjointement avec l'audio associé, pendant que l'enfant répète et visualise sa propre articulation sur un deuxième écran. Les auteurs de l'étude émettent alors l'hypothèse que ce retour pourrait bénéficier aux populations malentendantes, à condition que le patient soit capable d'interpréter ce qu'il voit, cela nécessitant de pouvoir y consacrer suffisamment de séances.

Depuis cette première étude de cas, de nombreuses autres études ont été réalisées, majoritairement dans les dix dernières années.

Issues principalement de pays anglophones, elles peuvent se répartir en quatre domaines : la rééducation du /ɪ/ anglais, les troubles articulatoires, les troubles de la parole chez les personnes malentendantes et l'apprentissage d'une langue seconde. Nous nous focalisons ici sur les trois premiers sujets, le quatrième ne relevant pas de la rééducation orthophonique même si le travail réalisé est similaire, comme en témoignent les récents travaux de Pillot-Loiseau *et al.* (2014) ou Wu *et al.* (2015).



**Figure 4.** Exemple de positionnement d'une sonde échographique et de l'affichage de l'image obtenue (haut). Images ultrasonores de la langue dans le plan sagittal médian (position de *repos*, en bas à gauche, et lors d'un [k], en bas à droite; extrait de Hueber & Denby, 2009).

La rééducation du /ɪ/ anglais occupe une place importante parmi les demandes des patients chez les orthophonistes nord-américains. En effet, si ce /ɪ/ peut s'articuler de deux manières, une seule est considérée comme correcte par les orthophonistes de ces pays. Adler-Bock *et al.* (2007) s'intéressent à deux adolescents de 12 et 13 ans présentant un trouble articulatoire résiduel sur cette consonne. Après 13 séances, ces patients sont capables d'articuler correctement le /ɪ/ dans certains contextes, même si les performances en conversation ne présentent pas d'amélioration. Les auteurs suggèrent donc une plus grande durée de rééducation. Bernhardt *et al.* (2008) s'intéressent au même trouble pour un groupe de 13 enfants pour lesquels les méthodes traditionnelles de rééducation n'ont pas fonctionné. Chacun bénéficie de

3 à 4 séances de rééducation en l'espace de deux jours, pour un total de 2 à 3 heures de rééducation ; il en résulte une plus grande précision acquise par l'ensemble des participants sur le phonème isolé ou dans un mot, mais pas en parole spontanée. Comme Adler-Bock *et al.* (2007), les auteurs suggèrent d'augmenter le nombre de séances. Plusieurs études de cas ont aussi été réalisées sur le même sujet, comme celles de Preston et Leaman (2014) sur une patiente de 59 ans victime d'un AVC ayant entraîné une aphasie de Broca, mais aussi Byun *et al.* (2014) sur un enfant de 11 ans ou encore Cavin (2015) sur un homme de 22 ans, dont les seuls troubles articulatoires concernent ce phonème, avec à chaque fois des progrès observés.

Les troubles articulatoires concernent tous les enfants chez qui, après l'âge de 5 ans, subsistent des troubles phonologiques ou articulatoires. Modha *et al.* (2008), Preston *et al.* (2013 ; 2016) et Cleland *et al.* (2015) se sont intéressés à cette pathologie avec des approches différentes. Par exemple, Cleland *et al.* (2015) mènent des études de cas sur 7 enfants présentant des troubles sur des consonnes variées. Les progrès des enfants sont mesurés par des bilans orthophoniques réguliers. De plus, la sonde étant fixée sous la mâchoire, des mesures articulatoires peuvent être faites sur les images échographiques. Cette étude souligne l'efficacité de ce retour visuel sur 12 séances de rééducation, avec un maintien des performances plusieurs semaines après la fin des séances.

L'apprentissage de la parole chez des personnes malentendantes de naissance bénéficiant d'une aide auditive constitue un troisième domaine d'application. Différentes études ont été menées par exemple par Adler-Bock *et al.* (2007) ou Bernhardt *et al.* (2003). Gallagher (2013) propose une critique intéressante de la littérature sur le sujet, où elle valide l'apport positif de l'échographie pour cette population. Elle souligne cependant la nécessité de poursuivre ces travaux sur des populations plus larges (> 7 personnes) et sur une durée plus longue (> 10 séances).

## C – Résumé

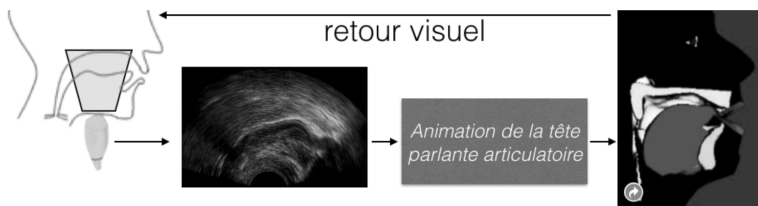
En résumé, la littérature reflète une grande diversité dans l'organisation et le contenu des séances de rééducation, ce qui rend difficile une éventuelle généralisation des résultats. Suivant les auteurs

et les troubles étudiés, plusieurs types d'organisation sont mis en œuvre : des sessions uniques, par exemple pour corriger un trouble articulaire très spécifique, ou des sessions réparties sur 20 semaines lorsque le trouble est plus important. La séance dure généralement entre 30 et 60 min, avec un temps consacré purement au retour visuel plus ou moins important. Les patients sont de tous âges, enfants, adolescents ou adultes, et se répartissent entre études de cas et études plus larges, certaines pouvant inclure jusqu'à 30 sujets. L'évaluation des performances se fait majoritairement par des bilans orthophoniques, même si des analyses acoustiques (enregistrements audio) ou articulatoires (tracés sur des images) complètent parfois ces bilans.

La diversité se retrouve aussi dans les approches adoptées. Par exemple, quand Bernhardt *et al.* (2008) alternent rééducation traditionnelle et retour visuel par échographie, Roxburgh *et al.* (2015) proposent une étude comparative entre le retour visuel échographique et l'illustration par une tête parlante articulaire. D'un autre côté, les travaux de Cleland *et al.* (2013) et Bacsfalvi *et al.* (2007) s'orientent vers la comparaison des deux technologies de retour visuel principales : l'EPG et l'échographie.

En conclusion, il est encourageant de constater que l'ensemble de la littérature converge vers le constat que l'utilisation de l'échographie a un impact positif sur la rééducation orthophonique, comme le soulignent par exemple Preston *et al.* (2013). Cleland *et al.* (2013) suggèrent toutefois qu'un retour visuel basé sur une tête parlante articulaire pilotée à partir d'images échographiques pourrait être le meilleur support à la rééducation orthophonique des troubles articulatoires.

Cet objectif est poursuivi par Fabre *et al.* (2014) qui proposent l'utilisation de techniques récentes d'apprentissage automatique (*machine learning* algorithme permettant l'apprentissage automatique d'une fonction ou d'un concept à partir d'exemples) pour animer automatiquement une tête parlante articulaire à partir d'images échographiques de la langue. Le système proposé est schématisé à la Figure 5. Des résultats prometteurs ouvrent des perspectives pour une utilisation en contexte clinique.



**Figure 5.** Schéma illustrant le retour articuloire.

## II – Apport de l’illustration visuelle échographique à la rééducation d’un trouble phonologique

Dans cette section, nous présentons les résultats d’une première étude visant à évaluer l’apport de l’illustration visuelle échographique à la rééducation d’un trouble phonologique. Notons que nous nous inscrivons dans le cadre du paradigme d’illustration et non de celui du retour. Notre démarche vise en premier lieu à évaluer concrètement la méthode la plus simple à mettre en place (c’est-à-dire l’illustration) et présentant l’intérêt le plus immédiat pour la communauté des orthophonistes.

L’étude s’appuie sur le logiciel UltraSpeech-player (Hueber, 2013) présenté ci-dessus et propose une visualisation de mouvements cibles acquis par échographie sur un locuteur de référence. Nous nous intéressons ici au trouble phonologique de substitution de [tʁ] par [kʁ], anomalie fréquemment rencontrée en orthophonie.

### A – Méthode

Quatorze enfants atteints d’un trouble phonologique avec substitution de [tʁ] par [kʁ], âgés de 5 ans 1 mois à 7 ans 5 mois, ont été répartis en deux groupes homogènes de sept participants : Groupe 1 (moyenne = 5 ans 11 mois), Groupe 2 (moyenne = 5 ans 9 mois).

Chacun a été suivi pendant trois semaines consécutives au rythme de 30 minutes hebdomadaires. Chaque séance consistait en un entraînement visant à développer la conscience proprioceptive et kinesthésique des articulateurs engagés dans la réalisation des sons [t] et [k] et du cluster [tʁ]. L’orthophoniste et l’enfant décrivaient d’abord les lieux articuloires des consonnes [t] et [k] pour que ce dernier puisse centrer son attention sur la position de sa langue lorsqu’on



lui demandait de produire ces consonnes. Un travail de répétition du groupe diconsonantique [tʁ] était ensuite proposé. L'association [tʁ], insérée dans une séquence V [tʁ] V, était scindée afin d'accéder à la prononciation correcte du phonème [t]. [t] et [ʁ] étaient donc séparés par une pause (*e.g.*, [at#ʁa]) initialement longue, puis réduite progressivement. Enfin, la dernière étape reposait sur la répétition du groupe diconsonantique [tʁ] suivi des voyelles [i], [a] et [u]. La vitesse de répétition était modulée grâce au curseur de ralentissement du logiciel, réglé pour des vitesses lente, moyennement lente et normale, correspondant aux débits de prononciation déjà utilisés par les enfants dans l'étape de répétition sans UltraSpeech-player.

Chaque enfant a bénéficié d'une phase d'entraînement avec l'illustration visuelle échographique rendue possible par le logiciel UltraSpeech-player (USp), et d'une autre phase d'entraînement sans illustration. L'ordre des phases avec et sans USp était contrebalancé entre les groupes. Comme l'illustre le Tableau 1, les performances des enfants étaient testées avant et après chaque intervention. Les tests de 12 minutes consistaient en des épreuves de production et d'identification programmées avec E-Prime 2.0 software (Psychology Software Tools, Inc., Sharpsburg, PA). La troisième séance, une semaine après le deuxième entraînement, n'impliquait que ces deux épreuves, afin d'évaluer la persistance des effets. Les performances ont été évaluées par enfant, pour chaque stade et pour chaque tâche (plus de détails sont disponibles dans Bach & Lambourion, 2014).

**Tableau 1.** Synthèse du programme expérimental d'évaluation des interventions orthophoniques avec et sans UltraSpeech-player (USp).

Etape	Séance 1			Séance 2			Séance 3
	T0	T1	T2	T2	T3	T3	T4
Groupe 1	Pré-test	Entraînement avec USp	Post-test	Pré-test	Entraînement sans USp	Post-test	Test
Groupe 2	Pré-test	Entraînement sans USp	Post-test	Pré-test	Entraînement avec USp	Post-test	Test

## 1 – Epreuve de production

Trente mots fréquents – 20 contenant le cluster [tʁ] et 10 contenant le cluster [ʁ] – ont été sélectionnés. Ces deux clusters étaient en position initiale du mot, en position intervocalique, ou encore finale. Les mots correspondaient à des référents concrets (*e.g.*, cra-

vate, trousse...) illustrés par 30 dessins réalisés pour l'étude. Après une phase de familiarisation avec le matériel, chaque image était présentée à l'écran et l'enfant devait prononcer le mot qu'il avait appris à lui associer. Ses productions acoustiques étaient enregistrées (H1 Accessory Pack APH-1 ZOOM). La prononciation de la séquence cible [tʁ] était cotée 0 lorsque qu'elle était bien produite et 1 pour toute erreur de production (substitution du [t] par [k] ou omission de [ʁ]). Ces cotations ont été effectuées par deux étudiantes en orthophonie, co-auteurs de l'article.

## 2 – Epreuve d'identification

L'autre épreuve présentait 18 logatomes enregistrés par une locutrice parlant un français sans accent marqué et contenant [t], [k], [tʁ] ou [kʁ], combinés dans des proportions identiques aux voyelles [e] [o] [ɑ̃]. Les consonnes étaient en position initiale (e.g., [tʁɑ̃]) ou intervocalique (e.g., [ɑ̃tʁɑ̃]). Après vérification de sa bonne connaissance des lettres T et K (et des phonèmes correspondants), l'enfant recevait un stimulus sonore via un casque audio et voyait apparaître à l'écran ces deux lettres côte à côte. Il devait appuyer sur la touche située sous la consonne entendue. Chaque son était présenté deux fois. Le logiciel E-prime saisissait automatiquement la cotation (0: réponse juste, 1: réponse fausse), et enregistrait les temps de réponse.

## B – Résultats

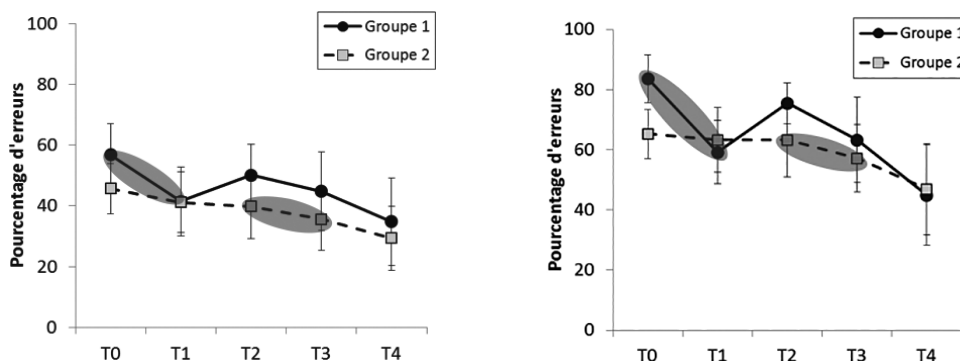
### 1 – Epreuve de production

L'évolution des pourcentages d'erreurs a été analysée aux différents stades de la prise en charge pour chaque groupe. En raison de la taille limitée de l'échantillon, les tests non paramétriques de Wilcoxon (*V*) et de Wilcoxon-Mann-Whitney (*W*) ont été utilisés pour les comparaisons intra- et inter-groupes respectivement.

Alors qu'à T0 le niveau de production ne différait pas entre les deux groupes ( $W = 31, p = .46$ ), le taux d'erreurs diminuait entre T0 et T1 pour le Groupe 1 entraîné avec USp ( $V = 27, p = .0313$ ) mais pas pour le Groupe 2 entraîné sans USp ( $V = 12, p = .28$ ) (Figure 6 gauche). Bien qu'ayant suivi un entraînement avec USp entre T2 et T3, le Groupe 2 n'améliorait pas significativement ses performances entre ces étapes ( $V = 19, p = .0935$ ). De manière générale, l'utilisation précoce d'USp dans la prise en charge s'accompagnait d'une

augmentation significative des performances entre T0 et T4 (Groupe 1,  $V = 21$ ,  $p = .0360$ ), suggérant le maintien du progrès une semaine après la fin de l'intervention. La progression entre T0 et T4 n'était par contre que marginale dans le groupe ayant bénéficié d'USp dans une phase tardive de la prise en charge ( $V = 20$ ,  $p = .0592$ ).

L'analyse détaillée des données montrait que l'amélioration des performances entre T0 et T1 dans le Groupe 1 avec USp s'expliquait surtout par une forte tendance au progrès pour produire [tʁ] en position initiale de mots ( $V = 25.5$ ,  $p = .058$ ), contrastant avec l'absence d'effet de l'entraînement sans USp à cette étape dans le Groupe 2 ( $V = 4$ ,  $p = .79$ ) (Figure 6 droite). Là encore, l'utilisation plus tardive d'USp, entre T2 et T3 (Groupe 2), n'a pas significativement modifié les performances pour [tʁ] en position initiale ( $V = 8.5$ ,  $p = .27$ ).

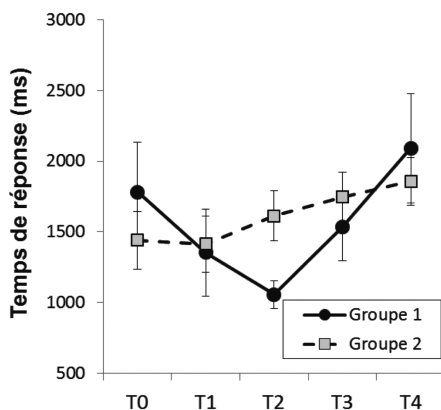


**Figure 6.** Pourcentages d'erreurs de production du cluster [tʁ] (dans les trois positions possibles) pour le Groupe 1 entraîné avec USp entre T0 et T1 et sans USp entre T2 et T3, et pour le Groupe 2 entraîné sans USp entre T0 et T1 et avec USp entre T2 et T3, sur l'ensemble des stimuli (gauche) ; idem pour les stimuli avec [tʁ] en position initiale (droite). Les barres représentent les erreurs-type. Les effets de l'entraînement avec USp sont mis en relief par les ellipses.

## 2 – Epreuve d'identification

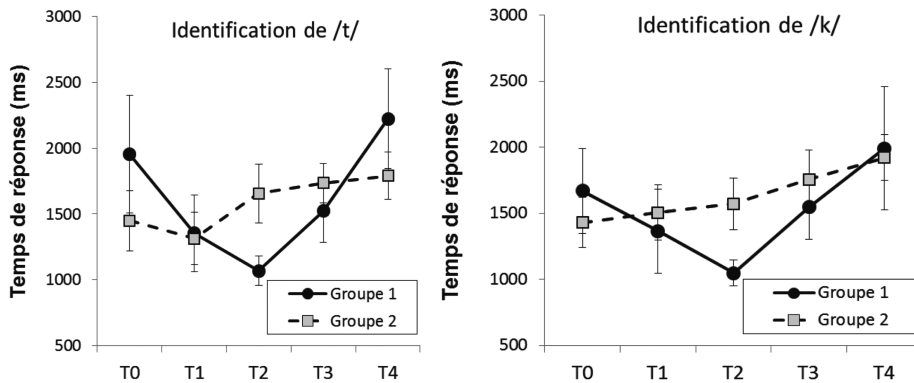
Les pourcentages d'erreurs et les temps de réponse moyens aux décisions exactes ont été traités avec des tests non paramétriques. Lorsque les résultats étaient regroupés pour [k] et [t], les pourcentages d'erreurs ne différaient pas entre les groupes à T0 ( $W = 18$ ,  $p = .44$ ) et aucune différence significative n'est apparue clairement au cours de la prise en charge. Il en était de même pour la vitesse de

réponse, qui ne différait pas entre les groupes à T0 sur l'ensemble des données ( $W = 26, p = .90$ ) comme pour les consonnes [k] ( $W = 26, p = .90$ ) et [t] ( $W = 25, p = .99$ ) considérées séparément, ce qui constituait une base intéressante pour les étapes ultérieures. Dans la suite de la prise en charge, lorsque les résultats pour [k] et [t] étaient rassemblés (Figure 7), l'unique effet significatif émergeait à T2 : le Groupe 1, qui avait été entraîné avec USp était alors plus rapide que le Groupe 2 qui n'avait pas encore été entraîné avec USp ( $W = 8, p = .0379$ ), sans que la qualité des réponses ne soit affectée.



**Figure 7.** Evolution des temps de réponse moyens d'identification consonantique pour les deux groupes d'enfants testés.

L'accélération des réponses du Groupe 1 après l'entraînement avec USp s'explique principalement par une accélération progressive des temps de réponse pour [t] entre T0 et T2 ( $V = 26, p = .0469$ ) (Figure 8 gauche) et par une accélération significative immédiatement après l'entraînement avec USp entre T0 et T1 pour [k] ( $V = 27, p = .0325$ ). Ce dernier progrès a permis à ce groupe de rester plus rapide que l'autre à T2 ( $W = 8, p = .0379$ ) (Figure 8 droite). Aucun progrès n'est par contre apparu entre T0 et T1 ou T2 dans le Groupe 2, entraîné sans USp. La comparaison inter-groupes des trajectoires d'évolution aux étapes plus tardives (T3, T4) a révélé des effets contradictoires entre vitesse et exactitude, délicats à interpréter, et sans doute dus à une stratégie favorisant la précision au détriment de la vitesse.



**Figure 8.** Evolution des temps de réponse moyens d'identification consonantique pour les deux groupes d'enfants testés pour [t] (à gauche) et [k] (à droite).

Les analyses des temps de réponse en fonction des structures CV, CCV et VCCV mettent en évidence des réponses plus rapides à T2 qu'à T0 pour le Groupe 1 pour [t] V ( $V = 28, p = .0156$ ), [k] V ( $V = 21, p = .0313$ ) et [kɤ] V ( $V = 26, p = .0469$ ), mais pas pour [tɤ] V ( $V = 16, p = .31$ ). Enfin, dans le contexte VCCV, les réponses étaient plus rapides à T2 pour le Groupe 1 qui venait d'être entraîné avec USp que pour le Groupe 2 entraîné sans USp, non seulement pour V [kɤ] V ( $W = 6, p = .0175$ ) mais aussi pour V [tɤ] V ( $W = 4, p = .0069$ ).

## C – Discussion

Les résultats de cette étude pilote montrent que l'illustration visuelle échographique des mouvements linguaux peut constituer un apport efficace à la prise en charge orthophonique d'un trouble phonologique chez de jeunes enfants. Alors que deux groupes d'enfants présentaient au départ une substitution de [tɤ] par [kɤ], une réduction significative de ce trouble est survenue, après une seule séance d'exercices ciblant la production de ce cluster, à condition que ces exercices s'accompagnent du retour visuel présenté par USp. Cela constitue un premier argument pour l'utilité de cet outil, mais aussi pour l'intérêt de cibler un ajustement de la conscience articulatoire des enfants, en tablant sur leurs capacités latentes de *lecture linguale*, pour améliorer à terme la production de la parole.

Cette possibilité d'améliorer la production de la parole par l'enrichissement des modalités de l'information fournie au locuteur sur ce qu'on attend de lui suppose la participation d'une représentation interne associant un résultat acoustique à une représentation des mouvements articulatoires basée sur la proprioception qui, lorsqu'elle est insuffisamment précise, peut être palliée par les apports d'une autre modalité, notamment visuelle. Des déficits très discrets de l'audition ou de la proprioception pourraient être à l'origine du trouble phonologique étudié ici, et il semble possible d'y suppléer en faisant appel à d'autres types d'informations pour établir une représentation plus adaptée: non seulement une description verbale des mouvements attendus, mais aussi une information visuelle analogique aidant à élaborer une image mentale visuelle réactivable pour les productions ultérieures.

Du point de vue clinique, les résultats enseignent aussi deux choses. Tout d'abord, les progrès suscités par l'utilisation d'USp pendant les exercices de rééducation sont surtout observés lorsque le cluster [tʁ] est en attaque de stimulus. Cela peut permettre au praticien d'anticiper des améliorations avant tout pour les mots contenant [tʁ] en position initiale. Une extension de l'étude pourrait renseigner en complément sur la chronologie des éventuels progrès pour ce cluster dans d'autres positions du mot.

Par ailleurs, du point de vue de la démarche clinique, les résultats renseignent sur le contexte le plus approprié pour introduire cet outil dans une rééducation. La petite taille de l'échantillon incite à la prudence, mais il apparaît qu'une utilisation du retour visuel dès l'instauration de la rééducation du défaut de production de [tʁ] est particulièrement efficace. Une interprétation possible est que l'entraînement lors d'une première séance ciblée sur [tʁ] sans illustration visuelle conduit l'enfant à reproduire son erreur dans un cadre assez formel qui détourne peut-être son attention de l'observation de ses sensations internes. L'absence d'information suffisamment compréhensible de façon directe n'incite sans doute pas non plus l'enfant à se détacher d'une représentation largement déviante. Ainsi renforcée par les exercices sans retour visuel, cette représentation inadéquate pourrait constituer un obstacle au changement proposé ensuite par l'apport d'une information d'un type nouveau par USp.

Par ailleurs, l'information apportée par USp revêt une forme inattendue pour l'enfant, ce qui l'interpelle sans doute suffisamment

pour mobiliser fortement son attention. Cette situation crée des conditions optimales pour intégrer les informations multiples offertes simultanément par différents biais (description de la part de l'orthophoniste, signal auditif, aspects proprioceptifs sollicités par les demandes de l'orthophoniste et indices visuels). Une telle intégration est sans doute plus facile à réaliser directement au moment de l'instauration des exercices intensifs sur l'unité de production posant problème, plutôt qu'après avoir involontairement renforcé par la répétition, dans une séance revêtant toute l'importance d'une situation thérapeutique avec un orthophoniste, une relation inadaptée entre une représentation phonologique, un son et un geste inadéquat. Lorsque le trouble de l'enfant est très ciblé, comme c'est le cas dans cette étude, les chances de succès de la rééducation semblent donc s'accroître si l'information apportée par USp est d'emblée associée à l'établissement d'une nouvelle représentation. L'efficacité de l'introduction précoce d'USp dans la rééducation est étayée par la significativité du progrès entre le pré-test (T1) et le dernier test (après les deux séances de prise en charge, à T4) seulement lorsqu'USp a été utilisé en première séance.

Enfin, les effets bénéfiques apportés par USp semblent s'établir via l'enrichissement d'une représentation suffisamment abstraite, sans doute phonologique, pour améliorer le traitement du phonème même en situation de perception, comme le montrent les progrès observés dans l'épreuve d'identification. Le fait que les progrès portent non seulement sur l'identification du phonème /t/, mais aussi sur l'identification du phonème /k/ produit correctement, suggère que c'est l'opposition entre les deux phonèmes qui est améliorée, et que l'illustration visuelle par le biais d'USp a permis d'accroître la précision du système phonologique.

## Conclusion et perspectives

Nous avons évoqué en introduction de cet article et dans notre revue de la littérature l'importance des canaux de retour sensoriels des actions motrices ainsi que la nécessité pour les sujets de maintenir ou de corriger des modèles internes de production et de perception de la parole cohérents. Nous avons présenté deux paradigmes



de représentation visuelle articulatoire pour la rééducation des troubles de l'articulation qui seraient susceptibles d'améliorer les capacités des patients en leur fournissant un complément d'information sur les phonèmes défectueux : l'illustration, et le retour articulatoire visuel proprement dit.

Nous avons réalisé une étude pilote visant à évaluer l'apport de l'illustration articulatoire à la correction d'un trouble phonologique de substitution pour un groupe d'enfants. Cette étude n'a pas d'équivalent en France, ni à l'échelle internationale, de par la combinaison d'au moins trois aspects : (1) le nombre relativement important d'enfants inclus par rapport aux autres études, (2) l'utilisation de films d'échographie linguale à vitesse normale ou ralentie pour l'illustration articulatoire pour des enfants, et (3) l'alternance systématique de séances de rééducation réalisées avec ou sans illustration visuelle, qui a permis d'utiliser chaque sujet comme son propre contrôle. Cette étude a montré qu'au-delà de son caractère attractif, l'illustration articulatoire semble apporter de réels bénéfices à la prise en charge des troubles phonologiques. En outre, cette approche présente l'avantage d'être très facile à utiliser dans la pratique d'un cabinet d'orthophonie, et même à domicile en autonomie, puisque UltraSpeech-player fonctionne avec un ordinateur ordinaire.

Dans le futur, nous nous proposons d'être beaucoup plus ambitieux, et d'évaluer l'apport d'un retour articulatoire visuel proprement dit, basé sur l'utilisation de l'échographie linguale auprès de deux populations : des adultes ayant subi une chirurgie bucco-pharyngienne comme dans l'étude de Blyth *et al.* (2016) et des enfants atteints de troubles articulatoires importants dus à des atteintes centrales. Des applications similaires chez le patient aphasique sont également en cours (Acher *et al.*, 2015).

## Remerciements

Ce travail a été financé par le programme de la Région Rhône-Alpes ARC6 *Technologies de l'Information et de la communication et usages Informatiques Innovants*, dans le cadre de la thèse doctorale de Diandra Fabre.

Acher, A., Fabre, D., Hueber, T., Amen, S., Lagarde, C., Badin, P., & Baciú, M. (2015). Retour visuel en rééducation orthophonique : étude d'un cas d'aphasie non-fluente. 6èmes Journées de Phonétique Clinique (JPC), Montpellier, France.

Adler-Bock, M., Bernhardt, B. M., Gick, B., & Bacsfalvi, P. (2007). The use of ultrasound in remediation of north American English /r/ in 2 adolescents. *American Journal of Speech-Language Pathology*, 16, 128 - 139.

Bach, C., & Lambourion, L. (2014). L'illustration visuelle échographique en orthophonie: un entraînement pour la prise en charge du trouble phonologique fonctionnel chez l'enfant. Mémoire présenté pour l'obtention du certificat de capacité d'orthophoniste (2012 - 2014).

Bacsfalvi, P., Bernhardt, B. M., & Gick, B. (2007). Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *Advances in Speech Language Pathology*, 9(1), 36-45.

Badin, P., Elisei, F., Bailly, G., Savariaux, C., Serrurier, A., & Tarabalka, Y. (2007). Têtes parlantes audiovisuelles virtuelles: données et modèles articulatoires - applications. *Revue de Laryngologie, Otologie, Rhinologie - European Review of ENT*, 128(5), 289-295.

Badin, P., Tarabalka, Y., Elisei, F., & Bailly, G. (2010). Can you 'read' tongue movements? Evaluation of the contribution of tongue display to speech understanding. *Speech Communication*, 52(6), 493-503.

Ballard, K. J., Smith, H. D., Paramatmuni, D., McCabe, P., Theodoros, D., & Murdoch, B. E. (2012). Amount of kinematic feedback affects learning of speech motor skills. *Motor control*, 16(1), 106-119.

Benoît, C., & Le Goff, B. (1998). Audio-visual speech synthesis from French text: Eight years of models, designs and evaluation at the ICP. *Speech Communication*, 26, 117-129.

Bernhardt, B. M., Bacsfalvi, P., Adler-Bock, M., Shimizu, R., Chenet, A., Giesbrecht, N., et al. (2008). Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clinical Linguistics & Phonetics*, 22(2), 149-162.

Bernhardt, B. M., Gick, B., Bacsfalvi, P., & Ashdown, J. (2003). Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics & Phonetics*, 17(3), 199-216.

Blyth, K. M., McCabe, P., Madill, C., & Ballard, K. J. (2016). Ultrasound visual feedback in articulation therapy following partial glossectomy. *Journal of Communication Disorders*, 61, 1-15.

Bothorel, A., Simon, P., Wioland, F., & Zerling, J.-P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Strasbourg: Institut de phonétique de Strasbourg.



Byun, T. M., Hitchcock, E. R., & Swartz, M. T. (2014). Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research*, 57(6), 2166-2130.

Cavin, M. (2015). The use of ultrasound biofeedback for improving English /r/. *Working Papers of the Linguistics Circle*, 25(1), 32-41.

Chen, Y.-P. P., Johnson, C., Lalbakhsh, P., Caelli, T., Deng, G., Tay, D., et al. (2016). Systematic review of virtual speech therapists for speech disorders. *Computer Speech & Language*, 37, 98-128.

Cleland, J., McCron, C., & Scobbie, J. M. (2013). Tongue reading: Comparing the interpretation of visual information from inside the mouth, from electro-palatographic and ultrasound displays of speech sounds. *Clinical Linguistics & Phonetics*, 27(4), 299-311.

Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics*, 29(8-10), 1-23.

Cowie, R., Douglas-Cowie, E., & Kerr, A. G. (1982). A study of speech deterioration in post-lingually deafened adults. *The Journal of Laryngology & Otology*, 96(02), 101-112.

Engwall, O. (2008). Can audio-visual instructions help learners improve their articulation? — An ultrasound study of short term changes. Proceedings of Interspeech 2008, Brisbane, Australia.

Epstein, M. A. (2005). Ultrasound and the IRB. *Clinical Linguistics & Phonetics*, 19(6-7), 567-572.

Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders*, XL, 481-492.

Eriksson, E., Bälter, O., Engwall, O., Öster, A.-M., & Kjellström, H. (2005). *Design recommendations for a computer-based speech training system based on end-user interviews*. Proceedings of the Tenth International Conference on Speech and Computers Patras, Greece.

Fabre, D., Hueber, T., & Badin, P. (2014). *Automatic animation of an articulatory tongue model from ultrasound images using Gaussian mixture regression*. Proceedings of Interspeech 2014, Singapur.

Fagel, S., & Madany, K. (2008). *A 3-D virtual head as a tool for speech therapy for children*. Proceedings of Interspeech 2008, Brisbane, Australia.

Gallagher, L. (2013). Critical Review: The effectiveness of ultrasound technology as a visual biofeedback tool on the productive speech intelligibility of adolescents and young adults with a hearing impairment. University of Western Ontario.



Hueber, T. (2013). *UltraSpeech-player: intuitive visualization of ultrasound articulatory data for speech therapy and pronunciation training*. Proceedings of Interspeech 2013, Lyon, France.

Hueber, T., Chollet, G., Denby, B., & Stone, M. (2008). *Acquisition of ultrasound, video and acoustic speech data for a silent-speech interface application*. Proceedings of the 8<sup>th</sup> International Seminar on Speech Production, ISSP8, Strasbourg, France.

Hueber, T., & Denby, B. (2009). Analyse du conduit vocal par imagerie ultrasonore. Dans A. Marchal & C. Cavé (Dir.), *L'imagerie médicale pour l'étude de la parole (Traité Cognition et Traitement de l'Information, IC2)* (pp. 147-174). Paris: Lavoisier.

Massaro, D. W., & Light, J. (2004). Using visible speech to train perception and production of speech for individuals with hearing loss. *Journal of Speech, Language, and Hearing Research*, 47, 304-320.

Matthies, M. L., Svirsky, M., Perkell, J., & Lane, H. (1996). Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant. *Journal of Speech, Language, and Hearing Research*, 39(5), 936-946.

Menin-Sicard, A., & Sicard, E. (2012). Intérêt de la visualisation de la position et du mouvement des articulateurs dans la prise en charge des troubles phonologiques. Paris: Les Entretiens de Bichat.

Mills, A. E. (1987). The development of phonology in the blind child. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: the psychology of lipreading* (pp. 145-161). London: Lawrence Erlbaum Associates.

Modha, G., Bernhardt, B. M., Church, R., & Bacsfalvi, P. (2008). Case study using ultrasound to treat /r/. *International Journal of Language & Communication Disorders*, 43(3), 323-329.

Montgomery, D. (1981). Do dyslexics have difficulty accessing articulatory information? *Psychological Research*, 43(2), 235-243.

Newell, K. M., Carlton, M. J., & Antoniou, A. (1990). The interaction of criterion and feedback information in learning a drawing task (English). *Journal of Motor Behavior*, 22(4), 536-552.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5), 382-407.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Perrier, P., Vick, J., et al. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28(3), 233-272.



Perkell, J. S., Matthies, M. L., Lane, H., Guenther, F. H., Wilhelms-Tricarico, R., Wozniak, J., et al. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, 22, 227-250.

Pillot-Loiseau, C., Kocjančič Antolík, T., & Kamiyama, T. (2014). Opposition /y/-/u/ chez 7 apprenantes japonophones avec et sans l'aide de l'ultrason lingual: 1. analyse acoustique comparative voyelles isolées /logatomes. Actes des 30<sup>èmes</sup> Journées d'Etude de la Parole, Le Mans, France.

Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology*, 22(4), 627-643.

Preston, J. L., & Leaman, M. (2014). Ultrasound visual feedback for acquired apraxia of speech: A case report. *Aphasiology*, 28(3), 278-295.

Preston, J. L., Maas, E., Whittle, J., Leece, M. C., & McCabe, P. (2016). Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. *Clinical Linguistics & Phonetics*, 30(3-5), 363-381.

Roxburgh, Z., Scobbie, J. M., & Cleland, J. (2015). *Articulation therapy for children with cleft palate using visual articulatory models and ultrasound bio-feedback*. Proceedings of the 18<sup>th</sup> International Congress of Phonetic Sciences, Glasgow, Scotland.

Shawker, T. H., & Sonies, B. C. (1985). Ultrasound biofeedback for speech training: Instrumentation and preliminary results. *Investigative Radiology*, 20(1), 90-93.

Stone, M., & Davis, E. P. (1995). A head and transducer support system for making ultrasound images of tongue/jaw movements. *The Journal of the Acoustical Society of America*, 98, 3107-3112.

Stone, M., & Shawker, T. H. (1986). An ultrasound examination of tongue movement during swallowing. *Dysphagia*, 1(2), 78-83.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212-215.

Turgeon, C., Prémont, A., Trudeau-Fisette, P., & Ménard, L. (2015). Exploring consequences of short- and long-term deafness on speech production: A lip-tube perturbation study. *Clinical Linguistics & Phonetics*, 29(5), 378-400.

Wu, Y., Gendrot, C., Hallé, P., & Adda-Decker, M. (2015). *On improving the pronunciation of French /r/ in Chinese learners by using real-time ultrasound visualization*. Proceedings of the International Congress of Phonetic Sciences, Glasgow, Scotland.



