

Étude du phénomène de convergence phonétique : comparaison enfants – adultes

Eva Goeseels^{1,2}, Bernard Harmegnies¹, Kathy Huet¹, Myriam Piccaluga¹, Virginie Roland¹, Clémence Verhaegen¹, Véronique Delvaux^{1,2}

(1) Institut de Recherche en Sciences et Technologies du Langage, Service de Métrologie et Sciences du Langage, Université de Mons, Belgique

(2) Fond National de la Recherche Scientifique, Belgique

eva.goeseels@umons.ac.be

RESUME

La convergence phonétique est définie comme la tendance au rapprochement des caractéristiques des productions de parole de locuteurs interagissant entre eux ou avec des dispositifs technologiques vocaux. Ce phénomène n'a été que peu étudié chez les enfants et se heurte à deux hypothèses concurrentes : les enfants convergeraient soit davantage soit moins que les adultes. Cette étude exploratoire traite donc de la convergence phonétique, en comparant des dyades enfants à des dyades adultes. 36 sujets (20 adultes et 16 enfants de 10-11 ans), ont produit les 75 phrases du corpus dans 5 conditions (4 individuelles et un jeu en duo) expérimentales successives. La fréquence fondamentale (f0) et le débit de parole moyen des énoncés ont été évalués. Les résultats indiquent que les deux groupes montrent de la convergence phonétique lors des tâches individuelles dans la mesure où ils se rapprochent de la cible, mais pour des mesures acoustiques différentes. Lors du jeu, la convergence phonétique a pu être mise en évidence dans les deux groupes, avec une grande variabilité inter-individuelle notamment en matière de dynamique temporelle.

ABSTRACT

Study of the phonetic convergence phenomenon: a comparison between children and adults.

Phonetic convergence is defined as the tendency, for speakers in an interaction, to change the characteristics of their speech production in order to match their interlocutor's, or to match a vocal technological device. This phenomenon hasn't been studied extensively with children and faces two competing hypotheses in this population: children should either be more convergent or less convergent than adults. This exploratory study focuses on phonetic convergence, by comparing dyads of children to dyads of adults. 36 participants (20 adults and 16 children aged 10-11) produced the 75 sentences of the corpus within 5 successive (4 individual and one duo game) experimental conditions. The pitch and the mean speech rate have been evaluated. The outcome of this study shows that the two groups were convergent for the individual conditions, but for different acoustic measures. During the game condition, phonetic convergence has been found in the two groups with, however, an important inter-individual variability, in particular in terms of a temporal dynamic.

MOTS-CLES : convergence phonétique, adaptation communicative, production, enfants

KEYWORDS: phonetic convergence, communicative adaptation, speech production, children

1 Introduction

La convergence phonétique est définie comme le « processus par lequel un locuteur a tendance à rendre son discours plus similaire à celui de son interlocuteur ». Il est donc question de rapprochement de caractéristiques vocales de locuteurs interagissant entre eux ou avec des dispositifs technologiques vocaux (Lelong, 2012). Ce phénomène d'adaptation communicative apparaît lorsque les locuteurs modifient les détails phonétiques fins de leur discours ou adoptent des aspects du répertoire phonétique/acoustique de leur interlocuteur pour lui ressembler davantage, même sans instructions à imiter. La convergence phonétique existe aussi dans sa forme inverse : la divergence phonétique. Celle-ci consiste, pour un locuteur, à augmenter les différences entre son discours et celui de son interlocuteur. L'étude de ce phénomène a débuté dans les années 70-80, avec les études de Giles et al. (ex : Giles et al., 1987) et leur théorie de l'adaptation communicative. Depuis lors, un grand nombre d'études a été réalisé, avec pour participants dans la plupart des cas des adultes.

Le phénomène de convergence phonétique a été mis en évidence à travers de nombreuses mesures acoustiques différentes, notamment la fréquence fondamentale (Fo), le débit de parole, les mesures de formants ou le VOT (*Voice Onset Time*) (entre autres : Babel and Bulatov, 2012 ; Babel, 2010 ; Kappes et al., 2009 ; Lelong & Bailly, 2011 ; Nielsen, 2011 ; Pardo et al., 2010, 2013 ; Sato et al, 2013 ; etc.). D'autres études recourent à une évaluation perceptive afin d'évaluer la convergence phonétique. Ces dernières recrutent alors des sujets n'ayant pas participé à la production, qui doivent juger perceptivement de la proximité phonétique entre différents sons de parole produits par les participants susceptibles de converger. Le paradigme mis en place pour étudier l'ampleur éventuelle de la convergence peut être de divers ordres : soit des tâches avec interaction directe, soit sans interaction directe. Dans les tâches avec interaction directe, les sujets sont amenés à interagir et discuter entre eux, dans celles sans interaction directe, les sujets n'interagissent pas directement entre eux, mais avec une « voix modèle », entendue via un casque par exemple. Au fil des études, les chercheurs ont pu mettre en évidence des facteurs pouvant influencer le processus de convergence. Notamment, le genre (ex : les femmes convergeraient davantage que les hommes (Namy et al., 2002)), le facteur social (ex : une personne ayant le rôle de leader dans une discussion pourrait moins converger vers son interlocuteur (Pardo, 2006)) ou encore le statut psycholinguistique des mots (ex : la fréquence lexicale (Goldinger, 1998)). Notons également que le phénomène de convergence phonétique est souvent cité pour soutenir le lien étroit entre perception et production (entre autres : Fowler et al., 2016 ; Pickering & Garrod, 2013 ; etc.) Chez les enfants, la convergence phonétique a été nettement moins étudiée (entre autres : Eaton & Ratner, 2013 ; Nielsen, 2011 ; etc.), mais il a été montré que les enfants étaient capables de converger, notamment pour des mesures de durées ou des caractéristiques phonétiques. Dans cette population, la capacité à ajuster ses productions de parole à des contraintes externes a été étudiée à travers le paradigme de perturbation/compensation de Ménard et al. (2016).

Nous avons décidé, dans le cadre de cette recherche exploratoire, de classifier et de séparer 3 concepts clés d'adaptation communicative, et qui seront utilisés dans ce travail de recherche. Tout d'abord, nous parlerons du concept de flexibilité, caractérisant un locuteur seul (sans interaction directe avec un autre locuteur), qui modifie les attributs phonétiques de sa parole en fonction d'un certain nombre de facteurs et de contraintes auxquelles il s'adapte. Ensuite, la convergence, décrira le phénomène de rapprochement des productions de parole entre deux locuteurs interagissant de façon directe. Enfin, nous utiliserons également de concept d'alignement, initialement décrit par Lelong en 2011, concept voisin de la convergence phonétique ayant lieu lorsque les deux sujets en interaction ne rapprochent pas leurs caractéristiques vocales, mais plutôt évoluent ensemble dans la même direction, de manière parallèle, sans pour autant converger.

La lecture de la littérature permet de mettre en évidence une double problématique. Premièrement, la convergence phonétique n'a été que très peu étudiée chez les enfants. Deuxièmement, l'examen de la littérature ne permet pas, *a priori* de trancher entre deux hypothèses concurrentes dans cette population. La première hypothèse stipule que les enfants, par rapport aux adultes, seraient davantage convergents (hypothèse « capacités imitatives »). En effet, les enfants sont en cours d'acquisition du langage et seraient donc plus flexibles en conversation que les adultes. L'acquisition du langage est en grande partie basée sur les capacités imitatives ainsi que sur l'environnement langagier dans lequel évolue l'enfant. L'article de Nguyen et Delvaux en 2016 traitant de l'imitation et son rôle à la fois dans le développement du langage chez l'enfant mais également dans le développement de la faculté de langage dans l'espèce humaine met en exergue que l'imitation entre individus est omniprésente dans le langage parlé, notamment durant l'enfance. La seconde hypothèse (hypothèse « immaturité ») soutient le postulat inverse, à savoir que les enfants seraient moins convergents, ou moins cohérents dans leur convergence. Effectivement, les enfants ne sont pas encore matures au niveau du contrôle moteur de la parole et ont encore beaucoup de variabilité non maîtrisée dans leurs productions ; ils n'auraient pas encore assez de représentations internes suffisamment précises (Barbier, 2016). Les enfants ne gèrent pas encore bien leur « instrument » et seraient plus « brouillons » dans leurs productions. Cette hypothèse a pu être mise en évidence suite aux résultats des études de Ménard et al. (2016 et précédentes) dans lesquelles il a été montré que les enfants compensaient une perturbation articulatoire via une stratégie essai-erreur, et non une correction graduelle des erreurs comme les adultes. Les enfants ne parviennent donc pas à stocker ces stratégies. Cela peut s'expliquer par l'immaturité de leur modèle interne des relations acoustico-articulatoires ainsi que du contrôle moteur de leur parole, ce dernier suivant un développement lent et progressif dont la maîtrise n'est semblable à celle de l'âge adulte qu'à la fin de l'adolescence (Smith et Zelanik, 2004).

L'objectif de cette étude est donc d'étudier les phénomènes de flexibilité (sans interaction directe) et de convergence (avec interaction directe) phonétique chez des enfants et des adultes. Nous nous sommes donc demandés si ces deux phénomènes étaient d'ampleur comparable dans les deux populations (pour la convergence phonétique, dyades d'enfants vs. dyades d'adultes). Nous nous sommes également demandés si ces phénomènes portaient sur les mêmes dimensions acoustiques de la parole dans les deux groupes (débit, f0).

2 Méthodologie

2.1 Participants et protocole

16 enfants (12 filles et 4 garçons) âgés de 10-11 ans, et 20 hommes adultes, de 27 ans et 5 mois en moyenne ont été recrutés. L'ensemble des sujets a pour langue maternelle le français et un développement typique. Les enfants ont été enregistrés dans une pièce calme de leur école ou leur école des devoirs à l'aide du dictaphone d'un *smartphone* (iPhone 8 ; fréquence d'échantillonnage : 48000 Hz). Les adultes ont été enregistrés dans les mêmes conditions, mais à leur domicile.

Le protocole est constitué de 5 tâches. La première tâche administrée aux sujets est une tâche de lecture, où ils ont dû lire les 75 phrases du corpus. Cette tâche permet d'avoir une ligne de base pour chaque sujet. La seconde tâche est une tâche de répétition dans laquelle les sujets entendent les 75

phrases du protocole prononcées par le locuteur modèle¹ via un casque, et ont pour instruction de répéter. L'objectif de cette tâche est d'évaluer dans quelle mesure les sujets suivent le locuteur modèle lors d'une consigne simple de répétition. Puis, les sujets ont passé une tâche d'imitation dans laquelle ils ont dû non seulement répéter mais ils ont eu comme consigne d'imiter ce qu'ils entendaient dans le casque. Cette tâche permet d'appréhender comment les sujets se comportent avec une consigne explicite d'imitation. Ensuite, une tâche d'inhibition a été administrée, où les sujets ont dû non seulement répéter les 75 phrases du protocole, mais ils ont eu pour consigne de s'empêcher d'imiter les voix qu'ils entendaient. Cette tâche est une tâche de plus haut niveau permettant de savoir si les sujets continuent d'imiter le locuteur modèle malgré eux. Pour ces 4 premières tâches du protocole, les sujets sont assis face à un ordinateur et travaillent avec une présentation PowerPoint. Pour cette partie, le sujet est seul avec l'expérimentateur, son binôme pour le jeu d'interaction n'est pas présent et n'entend pas ce qu'il se passe. Les tâches de répétition, imitation et inhibition sont des tâches sans interaction directe, où le phénomène étudié est la **flexibilité**. La 5e tâche du protocole est une tâche avec interaction directe, permettant d'étudier la **convergence** phonétique. Il s'agit d'une tâche interactive où les sujets sont à deux (assis face à face) et jouent à un jeu de survie. Ce jeu a été adapté de celui de l'étude de Kousidis et al. (2009). Dans cette tâche, les sujets doivent collaborer pour se mettre d'accord sur l'ordre d'importance dans lequel ils prendraient 15 objets pour leur survie en cas d'un naufrage fictif sur une île déserte.

Le corpus est constitué de quinze mots monosyllabiques représentant des objets appartenant au champ lexical de la survie, plage, île déserte, etc. Ils ont été sélectionnés via la base de données Lexique 3 (New & al. 2004) et appartiennent à trois niveaux contrastés d'âge d'acquisition et de fréquence lexicale. Les mots *torche*, *hutte*, *grill*, *bouée*, *bêche* sont peu fréquents, et acquis tard ; les mots *douche*, *table*, *pain*, *verre*, *sac* appartiennent au registre des mots fréquents, acquis tôt ; et les mots *piège*, *gaz*, *chaîne*, *couette*, *veste* sont de niveau intermédiaire. Ces mots ont été insérés dans une phrase porteuse avec une structure invariable se construisant comme ceci : « *Je mets le/la (...un des 15 mots...) en (...une position de 1 à 15...) position* ». Par exemple : « *je mets la torche en première position* ». Le mot-cible (ici : *torche*) se trouve au milieu de la phrase et non à la fin, évitant ainsi un risque d'hypo-articulation du mot-cible. Les locuteurs modèles ont donc produit ces 15 phrases au préalable de l'étude. Ensuite, nous avons pris la décision de manipuler le débit de ces 15 phrases, afin d'obtenir 5 réalisations de chaque phrase (du débit 1 (rapide) au débit 5 (lent)), soit un total de 75 phrases. Ces 75 phrases ont été proposées dans le même ordre pseudo-aléatoire au cours des 4 premières tâches du protocole. Les 5 niveaux de débits appliqués aux phrases des locuteurs modèles constituent le « débit cible » à atteindre pour les sujets.

2.2 Traitement et mesures

Les échantillons de parole recueillis auprès de chaque participant ont été analysés et traités à l'aide des logiciels SPPAS et Praat. Ensuite, l'ensemble des analyses statistiques a été réalisé avec IBM SPSS Statistics for Windows, version 25. SPPAS (Bigi, 2015) a permis, à partir de la transcription orthographique des enregistrements sonores, d'en proposer une transcription en Alphabet Phonétique International, et d'aligner cette transcription au fichier son correspondant. Le logiciel a également

¹ Deux locuteurs « modèles » ont été recrutés : un enfant (garçon, 9 ans) et un adulte (homme, 24 ans). Ils n'ont pas réalisé l'ensemble du protocole expérimental, mais ont simplement prêté leur voix qui a été enregistrée pour réaliser le protocole expérimental, dans la partie « flexibilité », sans interaction directe.

permis de segmenter le signal sonore en syllabes et ensuite en phonèmes. Les fichiers ainsi créés par SPPAS ont été exploités via Praat, afin de visualiser le résultat des segmentations et de procéder aux analyses acoustiques voulues.

Plusieurs mesures acoustiques ont été effectuées au moyen de scripts PRAAT personnalisés qui, via le paramétrage, permettait une adaptation au locuteur. Premièrement, la f0 et le débit ont été sélectionnés car il s'agirait, selon Pardo et al. (2010) et Sato et al. (2013), des deux paramètres acoustiques les plus sensibles à la convergence phonétique. Pour chaque énoncé, le débit et la f0 ont été mesurés. Le débit correspond au nombre de syllabes par secondes et la fréquence fondamentale a été calculée toutes les 5 ms et a ensuite été moyennée. A partir de ces mesures, des coefficients de flexibilité ($\frac{\text{Mesure} - \text{Ligne de base}}{\text{Cible} - \text{Ligne de base}} \times 100$) ont été calculés pour le débit et pour la f0. Il s'agit d'une mesure relative en pourcentage permettant d'apprécier l'ampleur du « chemin parcouru » par rapport au « chemin à parcourir » (écart entre la ligne de base du sujet et la production-cible). Cette mesure permet donc d'apprécier dans quelle mesure le locuteur étudié se rapproche ou s'éloigne du locuteur cible. S'il s'en rapproche (imitation), alors le coefficient aura une valeur positive. S'il s'en éloigne, le coefficient aura une valeur négative.

3 Résultats

3.1 Flexibilité

Les tâches sans interaction directe ont été exploitées grâce à des tests ANOVA à mesures répétées. La mesure répétée fait référence à la variable « condition », pour laquelle 3 niveaux ont été sélectionnés: répétition, imitation et inhibition. Ces analyses des parties individuelles du protocole permettent d'avoir une idée d'ensemble de la manière dont se comportent les sujets lors de conditions très contrôlées, méthodologiquement parlant, réalisées toujours sur les mêmes phrases, présentées dans le même ordre.

Dans cette partie, les analyses ont pu mettre en évidence qu'en général, quelle que soit la mesure acoustique étudiée, les sujets suivent bien les conditions. C'est-à-dire qu'ils parviennent à correctement répéter, imiter et inhiber les paramètres phonétiques de la parole du locuteur modèle, sauf dans les cas suivants:

- Les adultes ne parviennent pas à inhiber l'imitation de la f0 du locuteur modèle, en condition d'inhibition.
- Les enfants ne parviennent pas à inhiber l'imitation du débit cible imposé, quelle que soit la condition.

Concernant la f0, il existe un effet significatif ($F(2, 5182)=69,531$) de l'interaction entre le groupe et la condition sur le taux de flexibilité de la f0 des sujets. La Fig.1 illustre qu'en répétition, tous les sujets ont déjà un taux de flexibilité de 10 à 20% alors qu'il ne leur a pas encore été demandé d'imiter. En imitation, ce taux augmente pour les deux groupes. En inhibition, les enfants arrivent à suivre la consigne car leur taux de flexibilité diminue pour prendre une valeur négative (inhibition) ($M = -5.2 \%$, $SD = 3.3$), alors que pour les adultes, ce taux continue d'augmenter ($M = 31 \%$, $SD = 2.9$) par rapport à l'imitation ($M = 25.8 \%$, $SD = 2.1$). Ils n'arrivent donc pas à inhiber leur imitation de la f0 du locuteur modèle.

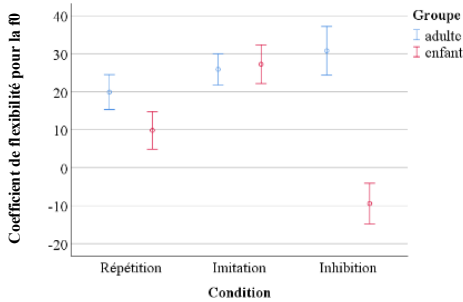


FIGURE 1: Coefficient de flexibilité de la f_0 (%) en fonction du groupe et de la condition

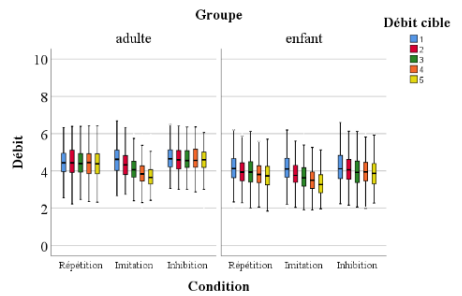


FIGURE 2: Débit (syll/s) en fonction de la condition, du groupe et du débit-cible

Concernant le débit de parole, on observe un effet significatif ($F(8, 5234)=2,666$) de l'interaction triple entre la condition, le groupe et le débit cible sur le débit des sujets. Comme illustré sur la Fig.2, les adultes suivent correctement la consigne en produisant ce qui est attendu d'eux: en condition de répétition, les adultes ne semblent pas changer leur débit, ils se contentent de simplement répéter et ne suivent donc pas le débit du locuteur modèle. En condition d'imitation, ils suivent le débit cible et en inhibition, ils arrivent à inhiber leur imitation du débit cible. Pour les enfants (à droite du graphique), cependant, nous pouvons voir qu'ils suivent le débit imposé, c'est-à-dire du plus lent au plus rapide, quelle que soit la condition. Ils imitent donc déjà le débit en répétition alors qu'il ne leur avait pas été demandé, ils imitent davantage en condition explicite d'imitation mais continuent d'imiter lorsqu'ils ont dû inhiber.

Les résultats ont également mis en exergue que la flexibilité est globalement plus importante pour la mesure acoustique "débit" que pour la f_0 , et que la comparaison des taux de flexibilité entre les enfants et les adultes diffère selon la mesure acoustique:

- Les adultes ($M = 25.533 \%$, $SD = 1.952$) ont une moyenne générale plus élevée que les enfants ($M = 14.365 \%$, $SD = 2.248$) au niveau du pourcentage de flexibilité pour la f_0 .
- Les enfants ($M = 47.698 \%$, $SD = 3.131$) ont une moyenne générale plus élevée que les adultes ($M = 36.981 \%$, $SD = 2.842$) au niveau du pourcentage de flexibilité pour le débit.

Un autre enseignement est que la tâche de répétition amène, pour tous les sujets, un taux de flexibilité positif. Ceci suggère donc que, quel que soit le groupe, les sujets ont déjà une tendance à la flexibilité phonétique en condition de répétition. Dans cette tâche, le taux de flexibilité s'élève, en moyenne, à 30% pour le débit, enfants et adultes confondus et à 15% pour la f_0 .

3.2 Convergence

Ensuite, les données récoltées lors du jeu ont été exploitées à l'aide d'analyses statistiques effectuées dyade par dyade, en comparant la condition de lecture (ligne de base) à la condition de jeu. Ces analyses ont permis de créer quatre catégories. Une catégorie de dyades convergentes, pour lesquelles l'analyse statistique de l'interaction entre la condition (comprenant 2 niveaux ici : lecture et jeu) et le locuteur (les deux locuteurs de la dyade) s'est révélée **significative** et le graphique représentant cette analyse permet de poser un diagnostic de convergence (les locuteurs se rapprochent pendant le jeu); une catégorie de dyades divergentes, pour lesquelles l'analyse statistique s'est révélée **significative** et le graphique représentant cette analyse permet de poser un diagnostic de divergence (les locuteurs

s'éloignent pendant le jeu); une catégorie de dyades montrant un pattern d'alignement, pour lesquelles l'analyse statistique s'est révélée **non significative** et le graphique représentant cette analyse permet de poser un diagnostic d'alignement (évolution parallèle dans la même direction); et une catégorie de dyades inclassables, pour lesquelles l'analyse statistique de l'interaction s'est révélée **non significative** et le graphique ne présentait pas d'alignement.

Par ailleurs, une analyse de l'évolution des productions des deux locuteurs de chaque dyade *au cours du jeu* a été effectuée. Sur la base des allures observées pour les droites de régression linéaire appliquées à ces données temporelles, chaque dyade a à nouveau été classée parmi 4 catégories (convergente, divergente, alignement et inclassable). A la suite des différents classements en catégories, un tableau permettant d'avoir une vue globale des résultats pour la partie convergence/jeu a été réalisé (Fig.3). Il existe une grande variabilité dans le phénomène en fonction de l'analyse effectuée, de la mesure acoustique étudiée ainsi que du groupe, mais la convergence n'est jamais absente des différents patterns de résultats, pour les adultes comme pour les enfants.

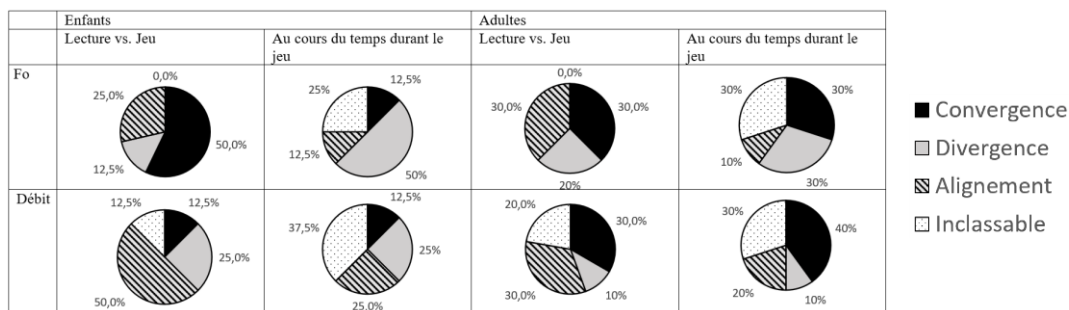


FIGURE 3 : tableau récapitulatif des différentes analyses de la partie « convergence »

4 Discussion

Dans cette étude, la convergence/flexibilité phonétique a été investiguée, en comparant ce phénomène chez des enfants et des adultes afin d'élargir le champ des recherches dans le domaine ainsi qu'amener des pistes de réponses à deux hypothèses quant à ce phénomène chez les enfants.

La flexibilité phonétique a pu être mise en évidence chez les enfants et les adultes mais les taux mesurés diffèrent selon les mesures acoustiques considérées. Les adultes se rapprochent plus de la f0 du modèle, à tel point qu'ils ne parviennent pas à inhiber le phénomène à la demande. Les enfants sont plus « flexibles » pour le débit, car ils suivent fidèlement le débit du locuteur-modèle quelle que soit la consigne. Dans ce cas, il serait éventuellement plus judicieux de ne pas parler de flexibilité phonétique mais bien de processus très automatisé, dont les enfants ont du mal de se défaire, quelle que soit la consigne. Ceci signifie que les enfants de 10-11 ans de cette étude sont incapables d'inhiber leur imitation, et qu'un processus très automatisé d'imitation implicite de la parole semble être en place. Cette découverte est à mettre en relation avec un fort lien entre perception et production, qui semble ici plus important chez les enfants que chez les adultes, en ce qui concerne le débit de parole. Dans la littérature, il a été avancé que la convergence phonétique reflétait un lien étroit entre perception et production, voire une intégration automatique entre perception et production (Pardo et al., 2018). La perception produirait très rapidement des cibles pour la production (Fowler et al. 2003), dont les enfants arriveraient difficilement à se défaire dans ce cas-ci. La perception spécifierait donc

les actions gestuelles pour la production, ce qui produirait une réponse imitative, impossible à inhiber pour les enfants. Par ailleurs, il a été mis en exergue que la mesure acoustique de débit amenait à plus de flexibilité pour tous les sujets que celle de f_0 . Cependant, seul le débit a été modifié artificiellement afin d'en obtenir 5 niveaux différents dans les stimuli. Ces résultats sont donc à prendre avec précaution.

Les analyses de la convergence phonétique ont montré des résultats variables selon la mesure étudiée, le groupe ou encore le type d'analyse réalisé. Ces résultats ne permettent pas de statuer sur une des deux hypothèses concurrentes énoncées avant le début de la recherche expérimentale, mais permettent de mettre en évidence que les enfants de 10-11 ans, comme les adultes, sont sensibles à la convergence phonétique. Ensuite, il est possible que la variabilité dans les résultats soit due à une certaine limite méthodologique dans la tâche utilisée pour évaluer la convergence phonétique avec interaction directe dans cette population. La littérature suggère que l'étude de la convergence phonétique en interaction chez les enfants serait plus fastidieuse car leur discours est plus disfluent mais également dû à une certaine timidité pouvant être plus importante dans ce type de tâche interactive que chez certains adultes (Coulston et al., 2002). Il est également possible que ces variabilités viennent du fait que les enfants procèdent, au niveau de leur adaptation communicative, à un mécanisme d'essai-erreur (Ménard et al., 2016). Le jeu ayant duré environ 5 minutes, les enfants n'auraient pas eu suffisamment de temps pour trouver leur stratégie d'adaptation communicative assez rapidement, dû au fait qu'ils procèdent par essai-erreur.

5 Conclusion

Cette étude a mis en évidence que les deux groupes montrent une tendance à la flexibilité phonétique et à la convergence phonétique. Cependant, les deux groupes montrent de la flexibilité et de la convergence pour des mesures acoustiques différentes, et différentes en fonction de la tâche réalisée ; une grande variabilité intra-groupe est également observée. Ce type de recherche expérimentale pose des défis notamment au niveau du choix des tâches pour étudier le phénomène chez les enfants. Ces résultats permettent de participer et de contribuer aux travaux précédents, permettant la compréhension du développement de la production de la parole chez les enfants et de ses mécanismes sous-jacents de contrôle moteur de la parole durant les interactions. L'étude de la convergence phonétique permet de confirmer davantage le lien production-perception, la production de la parole s'appuyant en permanence sur sa perception dans l'environnement, qui façonne à son tour la production. De plus, une meilleure compréhension des ressources en termes de flexibilité phonétique, tant chez les enfants que chez les adultes, pourraient constituer un marche-pied sur lequel appuyer des stratégies de compensation lors de la prise en charge de troubles de la parole.

Références

- BABEL M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, 39, 437-456.
- BABEL M., BULATOV D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech* 55, 231-248.
- BARBIER G. (2016). Contrôle de la production de la parole chez l'enfant de 4 ans : l'anticipation comme indice de maturité motrice.
- BIGI B. (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. *In the Phonetician - International Society of Phonetic Sciences*, ISSN 0741-6164, 111-112, I-II, 54-69.

- BOERSMA P., WEENINK D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.12, retrieved from <http://www.praat.org/>
- COULSTON R., OVIATT S., DARVES C. (2002). Amplitude convergence in children's conversational speech with animated personas.
- EATON C., RATNER N. (2013). Rate and Phonological Variation in Preschool Children: Effects of Modeling and Directed Influence. *Journal of speech, language, and hearing research* : JSLHR, 56.
- FOWLER C. A., SHANKWEILER D., STUDDERT-KENNEDY M. (2016). "Perception of the speech code" revisited: Speech is alphabetic after all. *Psychological Review*, 123, 125-150.
- GILES H., MULAC A., BRADAC J., JOHNSON P. (1987). Speech Accommodation Theory: The First Decade and Beyond, *Annals of the International Communication Association*, 10:1, 13-48.
- GOLDINGER S. D. (1998). "Echoes of echoes? An episodic theory of lexical access." *Psychological Review*, 105, 251-279
- KAPPE J., BAUMGAERTNER A., PESCHKE C., ZIEGLER W. (2009). Unintended imitation in nonword repetition. *Brain Lang*, 111, 140–151.
- KOUSIDIS S., DORRAN D., MCDONNELL C., E. COYLE. (2009). Time Series Analysis of Acoustic Feature Convergence in Human Dialogues. *SPECOM*. St Petersburg, Russian Federation.
- LELONG A., BAILLY G. (2011). Study of the phenomenon of phonetic convergence thanks to speech dominoes. In Esposito, Analysis of Verbal and Nonverbal Communication and Enactment. *The Processing Issues* (pp. 273-286). Springer, Berlin.
- LELONG A. (2012). Phonetic convergence in interaction. Université de Grenoble.
- MÉNARD L., PERRIER P., AUBIN J. (2016). Compensation for a lip tube perturbation in 4 year olds: Articulatory, acoustic, and perceptual data analyzed in comparison with adults. *The-Journal-of-the-Acoustical-Society-of-America*, 139(5), 25142531.
- NAMY L. L., NYGAARD L. C., SAUERTEIG D. (2002). "Gender differences in vocal accommodation: The role of perception." *Journal of Language and Social Psychology* 21, 422– 432.
- NEW B., PALLIER C., BRYSBART M., FERRAND L. (2021, 8 février) Lexique 3.83.
- NGUYEN N., DELVAUX V. (2016) Role of imitation in the emergence of phonological systems. *Journal of Phonetics*, Elsevier, 2016, 53, 46-54.
- NIELSEN K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39, 132-142.
- PARDO J. S. (2006). "On phonetic convergence during conversational interaction." *Journal of the Acoustical Association of America*, 119(4), 2382–2393.
- PARDO J. S., CAJORI JAY I., KRAUSS R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics*, 72(8), 2254-2264.
- PARDO J. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology*, 4(559).
- PICKERING M. J., GARROD S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 49–64.
- SATO M., GRABSKI K., GARNIER M., GRANJON L., SCHWARTZ J. L., NGUYEN N. (2013). Converging toward a common speech code: imitative and perceptuo-motor recalibration processes in speech production. *Frontiers in psychology*, 4, Article 422.
- SMITH A., ZELAZNIK H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental-Psychobiology*, 45(1), 22-33.