

Le développement de la coarticulation labiale anticipante en français : une étude préliminaire

Lucie Ménard^a, Marie-Agnès Cathiard^b et Christophe Savariaux^b

^aDépartement de linguistique et de didactique des langues

Université du Québec à Montréal, Canada

Tél.: ++(514) 987-3000, poste 3515 - Fax: ++(514) 987-4652

^bInstitut de la Communication Parlée

CNRS-INPG-Université Stendhal, Grenoble, France

Tél.: ++33 (0)4 76 82 43 37 - Fax: ++33 (0)4 76 82 43 35

Mél : menard.lucie@uqam.ca, cathiard@icp.inpg.fr, savariaux@icp.inpg.fr

ABSTRACT

This study investigates the development of lip rounding anticipatory coarticulation in [i(CCC)y] sequences, in French. Two 4-year-old children and two adult females served as subjects. Lip gestures were video-recorded and monitored using the ICP tracking system, and the internal lip area was computed. Various measurements in terms of lip area, velocity (first derivative) and acceleration (second derivative) were used to describe the anticipatory movement, in the framework of the *Movement Expansion Model* (MEM).

1. INTRODUCTION

Le développement de la parole nécessite le contrôle graduel des organes articulatoires, contrôle qui n'atteint la maturité qu'au terme de la première décennie de vie (Kent [1]). L'une des étapes marquantes de ce processus est la production de patrons coarticulatoires semblables à l'adulte, et plus spécifiquement des phénomènes d'anticipation. Un certain nombre d'études développementales ont mesuré l'ampleur de ce type de coarticulation, au plan acoustique, en tant qu'indice de maturité du contrôle moteur.

L'objectif de cette étude est de caractériser le développement de la coarticulation labiale anticipante dans des séquences [i(CCC)y], chez deux enfants de 4 ans et deux adultes français. Les données géométriques d'aire aux lèvres ont été analysées afin de décrire les différentes phases du mouvement de constriction, dans le cadre du *Movement Expansion Model* (MEM), auparavant décrit pour des adultes (Abry et Lallouache [2], Abry *et al.* [3]).

2. LES MODÈLES D'ANTICIPATION LABIALE

La production des voyelles arrondies se caractérise essentiellement par les deux composantes labiales de projection des lèvres et de fermeture ou de constriction. Chez l'adulte, l'analyse du décours temporel du geste de

protrusion labiale a donné lieu à quatre principaux modèles de l'anticipation, dans des séquences /iCy/ (voir l'évolution des approches de Benguerel et Cowan [4] jusqu'à Perkell et Matthies [5]). Le modèle « time-locked » suppose que l'arrondissement des lèvres débute à date fixe par rapport au début acoustique de la voyelle arrondie /y/. Le nombre de consonnes intervocaliques n'influence donc pas le début du mouvement. Au contraire, selon les tenants du modèle dit « look-ahead », le geste d'arrondissement débiterait dès la fin de la réalisation acoustique du /i/. L'intervalle temporel entre le début de ce geste et la réalisation acoustique du /y/ est donc expansible, c'est-à-dire variable en fonction du nombre de consonnes. Afin de rendre compte de données qui contredisent les deux modèles, le modèle hybride a été proposé, modèle qui prédit un mouvement lent de protrusion qui débute dès la fin de la voyelle non arrondie /i/, suivi d'un mouvement plus rapide dont la réalisation s'amorce à date fixe par rapport au début acoustique de la voyelle arrondie /y/. Enfin, selon le MEM (Abry *et al.* [3]), établi pour le français, la durée du mouvement de protrusion labiale de /i/ à /y/ est constante pour une séquence sans consonne intervocalique. Cette durée est appelée « constante d'exécution ». Elle augmente ensuite en fonction du nombre de consonnes intervocaliques, selon un coefficient d'expansion propre à chaque locuteur. D'abord développé à partir des mesures de protrusion labiale, le MEM a été adapté à des données d'aires aux lèvres (Abry *et al.* [3]).

Chez l'enfant, bien peu d'études ont tenté de modéliser l'anticipation des gestes articulatoires d'arrondissement labial. Signalons cependant Abelin *et al.* [6], qui ont observé six enfants (deux de 10 ans, trois de 8 ans et un de 7 ans) et un adulte, lors de la réalisation de suites $V_1C_nV_2$ où V_1 est non arrondie, V_2 est arrondie et C_n correspond à un groupe de consonnes non labialisées (de 1 à 7). Il appert que les adultes réalisent des mouvements anticipatoires de type « look ahead », alors que les enfants adoptent des stratégies qui se conforment plus au modèle « time-locked ». Une analyse attentive des

données présentées par les auteurs nous amène cependant à proposer une interprétation différente. Un enfant de 8 ans a des réalisations clairement de type « look-ahead », un autre (de 8 ans également) clairement « time-locked » tandis que les autres enfants ont des réalisations ne pouvant être classées selon l'un ou l'autre des modèles.

3. MÉTHODOLOGIE

3.1. Locuteurs et corpus

Afin de décrire les différentes phases du mouvement anticipatoire, deux enfants de 4 ans et deux adultes ont été enregistrés à l'aide du « poste visage-parole » développé à l'Institut de la Communication Parlée (Lallouache [7]). Les quatre sujets, de sexe féminin, étaient d'origine française. Le corpus consistait en un ensemble de 4 séquences de type [i(CCC)y], où C correspond à une suite de 0 à 3 consonnes non labialisées ([iy], [isy], [iky], [iksy], [iksty]). Pour les deux enfants, 3 à 6 répétitions de chacune des séquences ont été obtenues, alors que les locuteurs adultes ont produit 7 répétitions de chacune des formes. Les séquences cibles étaient insérées dans des phrases porteuses du type « La girafe Ikstu est rouge », produites en ordre aléatoire.

3.2. Procédure expérimentale

Les lèvres du locuteur, peintes en bleu, ont été filmées par une caméra de face, à un taux d'échantillonnage de 50 images par seconde (25 images entrelacées). Le signal audio a été numérisé à un taux d'échantillonnage de 22050 Hz. À l'aide d'un système de détection automatique des contours labiaux internes et externes développé à l'ICP, la valeur d'aire aux lèvres a été

déterminée pour chaque image. Une inspection manuelle des images a permis de corriger les erreurs de détection.

Le signal d'aire aux lèvres ainsi obtenu a été filtré par un filtre passe-bas de 7 Hz de fréquence de coupure, et les dérivées première et seconde, correspondant respectivement à la vitesse et à l'accélération, ont été générées. Conformément aux études antérieures, les événements suivants ont été étiquetés, sur les signaux synchrones de l'aire, de la vitesse et de l'accélération : les valeurs minimales et maximales d'aire aux lèvres ainsi que les valeurs maximales de vitesse et d'accélération. Sur le signal acoustique, la fin de la voyelle [i] a été identifiée par la disparition de la structure formantique sur le spectrogramme, alors que le début acoustique de [y] a été déterminé par l'apparition des formants. La différence, en secondes, entre le début de [y] et la fin de [i] correspond à l'intervalle d'obstruence (IO). Enfin, en considérant l'écart entre les valeurs d'aire aux lèvres, celle maximale pour [i] (événement 1 sur la figure 1) et minimale pour [y] (événement 4), trois événements supplémentaires ont été identifiés (Abry *et al.* [3]) : le temps correspondant à 10% de la diminution d'aire de [i] à [y] (événement 2), le temps correspondant à 90% de cette écart (événement 3) et le temps correspondant à une remontée de 10% de la valeur d'aire, suivant la valeur minimale de [y] (événement 5). L'écart entre les événements 2 et 3 est appelé « temps de chute » ou « time-falling » (TF) et l'écart entre 3 et 5 correspond à la « tenue » ou « hold » (H). La durée totale du mouvement est obtenue par la différence entre 2 et 5.

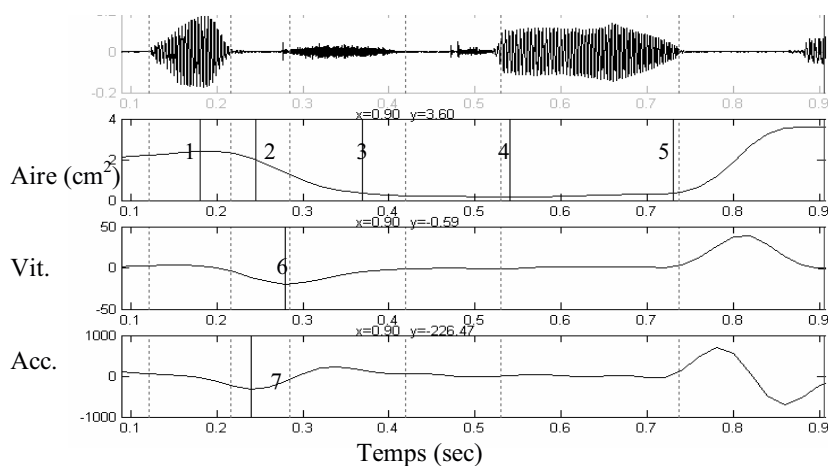


Figure 1 : Signal acoustique, décours de l'aire aux lèvres, de la vitesse et de l'accélération pour la séquence [iksty]. Identification des événements : 1= aire maximale de [i] ; 2=10% de la différence d'aire entre [i] et [y] ; 3=90% de la différence d'aire entre [i] et [y] ; 4=aire minimale de [y] ; 5=10% de la différence d'aire entre [i] et [y] suivant la valeur minimale de [y] ; 6=vitesse maximale (de signe négatif sur le graphique puisque l'aire diminue) ; 7=accélération maximale.

4. RÉSULTATS

4.1. Anticipation des événements de protrusion

Les valeurs de temps correspondant aux valeurs minimales et maximales identifiées sur les signaux d'aires, de vitesse et d'accélération ont été mises en relation avec la durée de l'IO. Chaque événement de protrusion est représenté pour chaque locuteur, à la Figure 2, en terme de pourcentages de IO. La formule suivante est utilisée :

$$(\text{évén.} - \text{fin acous. de [i]}) / \text{IO} * 100.$$

Une valeur de 0 signifie que l'événement correspond à la fin acoustique de [i] alors que la valeur de 100% signifie que l'événement correspond au début de [y]. Les deux locuteurs adultes sont identifiés par MA et CL, alors que les deux enfants sont représentés par les symboles AC_4 et EC_4. Comme le montre la figure 2, pour tous les locuteurs, la valeur minimale d'aire (Smin) est relativement peu anticipée par rapport à [y] (les valeurs avoisinent 100% ou sont supérieures à 100%). Le mouvement de constriction — tel que manifesté par la valeur d'anticipation de la vitesse maximale et de l'accélération maximale — est amorcé bien avant le début de [y]. Il peut être même commencé dès le début acoustique du [i] chez 3 locuteurs (sauf CL).

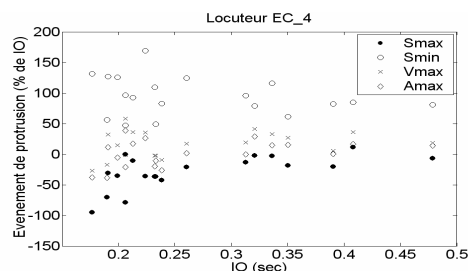
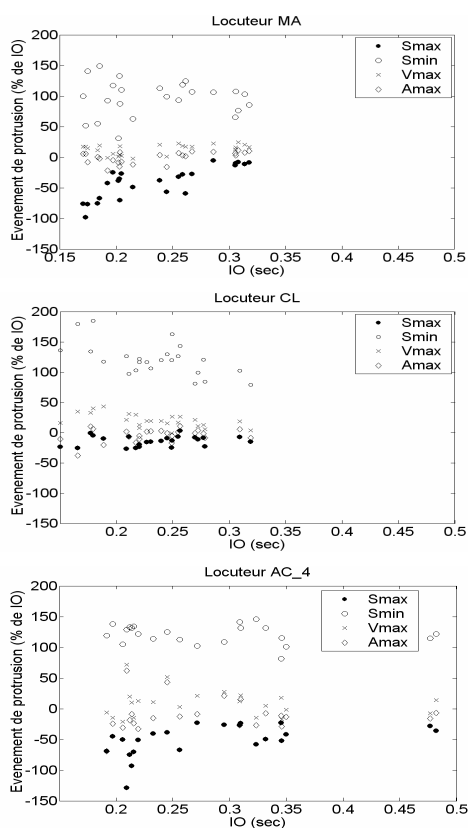
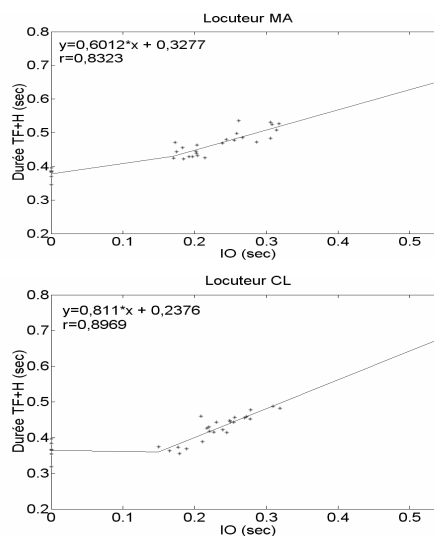


Figure 2 : Organisation temporelle des phases de l'anticipation de l'aire aux lèvres : aire maximale (Smax), aire minimale (Smin), vitesse maximale (Vmax) et accélération maximale (Amax). Le temps t d'apparition de chaque événement est exprimé en pourcentage par rapport à l'intervalle d'obstruence (voir texte).

4.2. Phases de constriction et IO

Afin de confronter les données au MEM, la durée totale du mouvement de constriction (TF+H) a été mise en relation avec la durée de IO, pour chaque locuteur, à la figure 3. Pour chaque figure, les valeurs pour lesquelles IO est nul correspondent aux séquences [iy] (constante d'exécution), alors que les durées croissantes de IO représentent les séquences [iCy], [iCCy] et [iCCCy]. Une analyse de régression linéaire a été effectuée sur les données, en excluant les séquences [iy]. La droite résultante, de même que son équation et son coefficient de corrélation r , figurent sur chaque graphique. La première section de la droite, de pente presque nulle, a été obtenue en reliant la moyenne de TF+H pour les séquences [iy] au premier point de la droite de régression. Rappelons qu'un modèle purement « time-locked » prédit une équation de pente nulle, alors qu'un modèle « look ahead » prédit une pente de valeur 1.



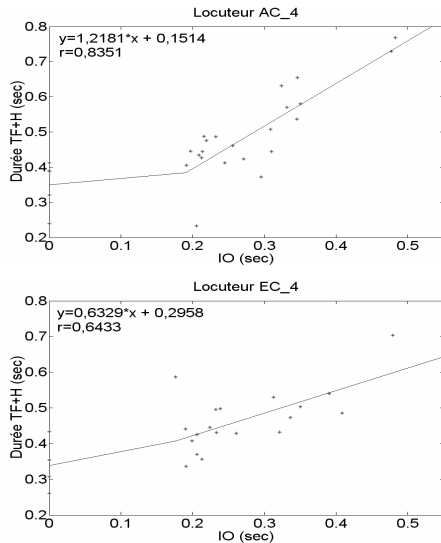


Figure 3 : Relation entre la durée totale du mouvement (durée de chute TF + tenue H) et la durée de l'intervalle d'obstruction IO, en secondes. Pour chaque graphique, un ajustement linéaire en deux parties a été effectué.

La figure 3 montre d'abord une stabilité remarquable de la durée du mouvement pour [iy] (constante d'exécution), d'un locuteur à l'autre (environ 350-380 ms). Il est également remarquable que le phénomène d'anticipation du mouvement, de même que son expansion en fonction du nombre de consonnes, est déjà présent chez nos locuteurs dès l'âge de 4 ans. Les coefficients de corrélation r , élevés pour nos 4 locuteurs, confirment la validité de la régression linéaire. Les valeurs plus élevées des constantes associées à nos locuteurs, par rapport aux données de Abry *et al.* [3], sont sans doute attribuables au caractère accentué du mot cible, suivi d'une pause. Ces conditions augmentent la durée de la tenue de [y]. Enfin, les valeurs de pente des deux locuteurs adultes (0,6012 et 0,811) et de l'enfant EC_4 (0,6329) ne s'approchent ni d'un modèle « time-locked », ni d'un modèle « look ahead ». Cependant, la valeur élevée de pente de la locutrice AC_4 (1,2181) contraste avec les données des adultes (Abry *et al.* [3]). Cette anticipation de type « look ahead » diverge également des résultats obtenus par Abelin *et al.* [6]. Il serait tentant de considérer ce comportement comme l'indice d'une fenêtre plus large de planification et/ou d'exécution du mouvement. De plus amples investigations sont actuellement en cours afin de confirmer cette hypothèse.

5. CONCLUSION

Les données présentées montrent la robustesse du MEM dans la représentation des phénomènes d'anticipation labiale lors de la production de séquences [i(CCC)y]. En effet, la valeur de la constante d'exécution telle que représentée par la suite

[iy] se situe aux environs de 350 ms, chez deux enfants de 4 ans et deux adultes. Notons qu'aucun des locuteurs ne produit des gestes articulatoires dont l'anticipation peut être représentée par les modèles « time-locked » ou « look-ahead ». Ces données peuvent par contre être bien prises en compte dans le cadre du MEM avec un coefficient d'expansion paramétrable par locuteur. Cette étude exploratoire sera suivie d'une étude plus approfondie de plusieurs sujets enfants. Des mesures longitudinales permettront également d'observer les éventuelles évolutions des coefficients d'expansion de chaque enfant, en fonction de l'âge.

Remerciements : Merci à Alain Arnal et Hélène Lœvenbruck pour leur aide technique et logicielle, et à Christian Abry pour ses conseils. Un merci tout spécial à nos petits locuteurs ainsi qu'à leurs parents !

BIBLIOGRAPHIE

- [1] R. D. Kent. Anatomical and Neuromuscular Maturation of the Speech Mechanism : Evidence from Acoustic Studies, *Journal of Speech and Hearing Research*, 19, pages 421-447, 1976.
- [2] C. Abry et M. T. Lallouache. Modeling lip constriction anticipatory behaviour for rounding in French with the MEM (Movement Expansion Model), *XIIIth Int. Congr. of Phonetic Sciences*, 4, pages 152-155, 1995.
- [3] C. Abry, M.-A. Cathiard, R. El Abed, M. T. Lallouache, M.-C. Leroy, P. Perrier, F. Poveda et C. Savariaux. Silent speech production : anticipatory behaviour for 2 out of the 3 main vowel gestures/features while pausing. *1st Esca Tutorial and Research Workshop on Speech Production Modeling*, pages 101-104, 1996.
- [4] A.P. Benguerel et H.A. Cowan. Coarticulation of upper lip protrusion in French. *Phonetica*, 30, pages 41-55, 1974.
- [5] J.S. Perkell et M. L. Mathies. Temporal measures of anticipatory labial coarticulation for the vowel /u/ : Within- and cross-subject variability. *Journal of the Acoustical Society of America*, 91 (5), pages 2911-2925, 1992.
- [6] A. Abelin, I. Landberg et L. Persson. A study of anticipatory labial coarticulation in the speech of children, *Phonetic Experimental Research at the Institute of Linguistics University of Stockholm*, 2, pages 2-18, 1980.
- [7] M. T. Lallouache. *Un poste « Visage-Parole » couleur. Acquisition et traitement automatique des contours des lèvres*. Thèse de l'INP, Grenoble, 1991.