

Développement de la parole et émergence de la structure prosodique chez l'enfant : une étude de l'accent d'emphase en français

LUCIE MÉNARD et MÉLANIE THIBEAULT
Université du Québec à Montréal

1. INTRODUCTION

Le développement de la parole requiert la maturation des différents mécanismes du conduit vocal. La maîtrise séquentielle des mouvements du larynx, de la mâchoire et de la langue, par exemple, influence considérablement le répertoire vocal du jeune enfant (Vihman 1996; MacNeilage 1998). Apprendre à parler signifie non seulement parvenir à produire les routines articulatoires correspondant aux unités phonologiques de la langue maternelle, mais également apprendre à intégrer une séquence de segments au sein d'une structure prosodique hiérarchisée (Beckman 1986; Vihman et al. 2006). Il a été montré que les locuteurs adultes signalent la structure prosodique de leur langue par des indices acoustiques et articulatoires réalisés par les articulateurs glottiques et supraglottiques (de façon indépendante ou conjointe)¹. En anglais et en français, par exemple, les syllabes produites sous emphase contrastive sont articulées avec une valeur plus élevée de fréquence fondamentale (F0) (Touati 1987), de durée et/ou d'intensité, par rapport à leur contrepartie non accentuée. Qui plus est, ces voyelles sous emphase sont aussi associées à des changements dans la dynamique des articulateurs supraglottiques (Beckman et al. 1992; Løvenbrück 1999; Erickson 2002). Chez les enfants qui ne montrent aucun trouble de développement phonologique, les mécanismes glottiques, ayant une influence prépondérante

Nous tenons à remercier Hélène Løvenbrück pour ses conseils judicieux et Caroline Émond pour son aide lors du traitement des données. Merci également aux évaluateurs anonymes pour leurs commentaires qui ont grandement amélioré la qualité de l'article. Ce travail a été réalisé grâce au soutien financier du Fonds québécois de recherche sur la société et la culture (FQRSC), le Conseil de recherche en sciences humaines du Canada (CRSH) et le Conseil de recherche en sciences naturelles et en génie du Canada (CRSNG).

¹Il est à noter cependant que les manifestations acoustiques observées ne peuvent être exclusivement reliées à un groupe d'articulateurs en particulier. Bien que, traditionnellement, les paramètres de fréquence fondamentale (F0) et d'intensité soient considérés comme étant contrôlés en grande partie au niveau laryngal (glottique), il n'en demeure pas moins que la tension exercée sur les cordes vocales par la hauteur de la langue affecte la valeur de F0 résultante (Esling 2005).

sur la fréquence fondamentale et l'intensité, sont contrôlés avant les mécanismes supra-glottiques au cours du développement (Vihman 1996). Ainsi, il est possible que différentes stratégies articulatoires soient impliquées dans la production de syllabes accentuées chez les adultes et les enfants. Dans une étude portant sur les corrélats géométriques de l'accent d'emphase chez des enfants de quatre ans et de huit ans et chez des adultes, Ménard et al. (2006) ont montré que les syllabes accentuées, bien que produites avec une augmentation marquée de F0 et d'intensité chez tous les sujets, ne sont pas produites par une plus grande ouverture des lèvres par rapport aux syllabes non accentuées chez les enfants, contrairement aux adultes. Les capacités réduites de déplacement des articulateurs chez l'enfant limitent donc l'inventaire des stratégies possibles afin de réaliser les différentes représentations prosodiques sous-jacentes. Alors que l'expérience décrite dans Ménard et al. (2006) portait sur un paramètre géométrique (l'aire aux lèvres), le présent article a pour but de décrire les déplacements des articulateurs labiaux et mandibulaires permettant de réaliser ce paramètre géométrique chez des enfants âgés de quatre et de cinq ans et chez des adultes.

2. L'EMPHASE CONTRASTIVE CHEZ LES ADULTES

L'emphase contrastive est définie comme une condition prosodique impliquant la sélection, par le locuteur, d'un constituant donné du message par opposition à un autre constituant dans une comparaison paradigmatique. L'emphase contrastive peut être réalisée syntaxiquement (le constituant cible est mis en relief par le biais d'une construction clivée, par exemple) ou prosodiquement (le constituant cible est mis en relief par des traits prosodiques, comme par exemple la fréquence fondamentale (F0), la durée et/ou l'intensité), ou de façon combinée par la syntaxe et la prosodie. L'expérience décrite dans le présent article implique l'emphase contrastive réalisée à la fois par la syntaxe et la prosodie. Les études menées en anglais ont montré que ce type de proéminence est associé à des mouvements labiaux plus grands, plus rapides et plus longs pour les voyelles /a/ et /i/ (Cho 2002, 2005). Pour la voyelle /a/, une position plus basse de la mâchoire a été observée, par rapport à la même voyelle réalisée en condition neutre (Erickson 1998). Pour le français, Lævenbruck (1999) mentionne également que les positions de la langue pour la syllabe [la] produite sous emphase contrastive sont plus basses dans la cavité buccale, par rapport à la condition neutre. Ces effets articulatoires de l'emphase contrastive sont souvent reliés à des changements spectraux. En effet, les voyelles ouvertes comme /a/ sont réalisées acoustiquement par une valeur de F1 plus élevée, ce qui confère à cette voyelle une position plus périphérique dans l'espace acoustique sous emphase (Cho 2005). Il est important de souligner que dans la plupart des études portant sur les corrélats de l'accentuation, une importante variabilité entre locuteurs est observée.

Cette étude porte sur l'emphase contrastive en français. Contrairement à l'anglais, le français est une langue à accent fixe, pour laquelle les accents primaires (*pitch accent*) affectent potentiellement la syllabe finale d'un mot de classe ouverte (Lacheret-Dujour et Beaugendre 1999). Ces accents coïncident ainsi avec la frontière droite d'un domaine prosodique comme le syntagme accentuel ou le syntagme

intonatif (Jun et Fougeron 2000). Les syllabes produites sous accent primaire en français sont généralement plus longues et associées à une valeur de fréquence fondamentale (F0) plus élevée que la syllabe précédente inaccentuée. Ce type d'accent diffère de l'accent d'emphase contrastive en ce que ce dernier affecte généralement la première syllabe d'un mot contenant deux syllabes, mais peut également affecter la seconde syllabe ou les deux syllabes (Di Cristo 1998; Jun et Fougeron 2000). Sous accent d'emphase, les syllabes sont produites avec une intensité accrue et une importante montée de F0 (Lacheret-Dujour et Beaugendre 1999; Jun et Fougeron 2002), par rapport à leurs contreparties inaccentuées. La séquence sous emphase est suivie d'un plateau bas de F0 (Touati 1987).

Dans le but d'expliquer les effets de la proéminence accentuelle sur les patrons cinématiques des articulateurs supraglottiques, deux modèles sont proposés. Selon une première approche, les effets de l'accent sur les articulateurs seraient reliés à un phénomène d'hyperarticulation locale, donnant lieu à une diminution de la coarticulation entre les segments des syllabes accentuées (De Jong 1995). Une approche similaire de l'accent d'emphase est également proposée par Lævenbruck (1999) pour les mouvements de la langue, dans une étude des stratégies utilisées par deux locuteurs pour produire des syllabes sous accent d'emphase en français. Selon un second modèle des corrélats articulatoires de l'accent, le locuteur produirait les syllabes sous emphase avec une position de mâchoire et de langue plus basse dans le but d'augmenter l'aire du conduit vocal, et ainsi d'augmenter l'amplitude de la voyelle. Ce modèle est dit *modèle d'expansion de sonorité* (Beckman et al. 1992). Pour les voyelles ouvertes, les deux modèles permettent de prédire des corrélats de l'accent d'emphase similaires, c'est-à-dire un abaissement de la mâchoire et de la langue et une distance interlabiale plus grande pour les voyelles produites sous emphase contrastive, par rapport aux mêmes voyelles produites en condition neutre.

3. LE DÉVELOPPEMENT DES CORRÉLATS DE L'ACCENTUATION

D'un point de vue développemental, puisque les mécanismes glottiques et supraglottiques impliqués dans la réalisation du contraste accentué vs. non accentué sont maîtrisés à différents stades de croissance (Vihman 1996), il est prédit que les enfants exploitent différentes stratégies articulatoires par rapport aux adultes pour signaler l'emphase contrastive. Par exemple, il est possible que la maîtrise des articulateurs supraglottiques par les jeunes enfants ne soit pas suffisamment fine (en termes de vitesse et/ou d'ampleur du déplacement) pour exploiter toute l'échelle de la dimension d'hypo-hyperarticulation, comme c'est le cas pour l'adulte. Dans ce cas, les corrélats articulatoires-acoustiques de l'accent d'emphase contrastive émergeraient graduellement au cours du développement de la parole.

La capacité d'utiliser la syntaxe et la prosodie afin de réaliser l'emphase contrastive a été attestée chez des enfants dès l'âge de trois ans (Hornby et Hass 1970). Cependant, les corrélats de la proéminence accentuelle ont été généralement étudiés pour les syllabes sous accent lexical en anglais. Les paramètres prosodiques traditionnels de F0, de durée et d'intensité sont manipulés très tôt. Par exemple, Connaghan et al. (2001) montrent que des enfants de deux ans et de trois ans et

six mois produisent des augmentations de F0 et d'intensité comparables à celles des adultes lors de la production de syllabes accentuées. Allen et Hawkins (1980) obtiennent des résultats différents : dans leur étude, l'ampleur des contrastes selon ces dimensions acoustiques est réduite pour des enfants de deux à quatre ans, par rapport à celle produite par des adultes. (Une revue détaillée d'études similaires se trouve dans Kehoe et al. 1995.) La capacité de réaliser des contours intonatifs comparables à ceux des adultes est observée tôt en français (Konopczynski 1986). En français également, il a été montré que les enfants âgés de deux ans réalisent des allongements syllabiques sous accent comparables à ceux produits par les adultes (Konopczynski 1986). Dans des études détaillées des paramètres de durée, d'intensité et de F0 produits sur des séquences disyllabiques en français et en anglais, entre autres, Vihman et al. (1998) et Vihman et al. (2006) montrent qu'au tout début de la période des premiers mots, les patrons rythmiques de chacune des langues ne sont pas reproduits par l'enfant. Les patrons rythmiques alors produits par les jeunes enfants sont davantage de type syllabique qu'accentuel et ce, peu importe leur langue maternelle. Ce stade correspond, en moyenne, au début de la première année de vie. Lorsque l'enfant atteint une étape de maîtrise lexicale plus importante (correspondant, en moyenne, à 1,5 an-2 ans), les patrons de F0, de durée et d'intensité produits correspondent beaucoup plus aux patrons produits par les adultes locuteurs de leur langue maternelle.

En revanche, le développement des capacités de déplacement des articulateurs supraglottiques n'est terminé qu'au terme de la première décennie de vie (Vihman 1996; Green et al. 2002). Bien que certains muscles (comme l'orbicularis oris, par exemple) soient recrutés très tôt dans des tâches réflexes (la tétée, par exemple), il n'est pas clair que ce comportement soit le prélude des mouvements volontairement contrôlés de la parole (lors de l'articulation d'une voyelle arrondie, par exemple). Ruark et Moore (1997) ont en effet démontré que les patrons d'activation musculaire différaient selon que la tâche accomplie en est une de parole ou non. Dans la plupart des études développementales, une diminution de la vitesse des articulateurs est observée chez les enfants (par exemple, Smith et Goffman 1998). En ce qui concerne la durée, les résultats suggèrent que les enfants produisent en moyenne des segments plus longs que ceux des adultes (Smith et Goffman 1998). Malgré cette réduction de vitesse, le déplacement des articulateurs chez les enfants est comparable à celui des adultes. Compte tenu du fait que la taille des articulateurs de la parole, chez les enfants, est plus petite que celle des adultes, les enfants semblent produire des mouvements relativement plus grands que ceux des adultes. Très peu d'études ont eu pour but d'étudier les corrélats articulatoires des proéminences accentuelles chez les enfants, à l'exception des travaux de Goffman et Malin (1999) et de Goffman (2004), qui visent à relier les facteurs moteurs et linguistiques dans le développement de la parole. Dans ces études portant sur l'anglais, il est mentionné que la production des patrons d'accentuation trochaïques (accent sur la première syllabe d'un mot disyllabique) sont mieux maîtrisés que les patrons d'accentuation iambiques (accent sur la seconde syllabe d'un mot disyllabique). La séquence rythmique produite par défaut par l'enfant serait de type trochaïque, comme le suggère le fait que les déplacements articulatoires sur les séquences de type CVcv sont plus grands sur la première syllabe, accentuée, et sont réduits sur la seconde syllabe, inaccentuée. Il en est autrement

des séquences iambiques, pour lesquelles les deux syllabes, pourtant porteuses de différents degrés d'accentuation, ne sont pas différenciées en ce qui a trait aux déplacements des articulateurs. Il est possible que ce biais relatif aux déplacements limités des articulateurs soit relié à la structure métrique de l'environnement linguistique, l'anglais étant une langue à structure principalement trochaïque. Toutefois, Vihman et al. (1998) ont démontré que vers la fin de la période du babillage, les enfants ayant le français comme langue maternelle produisent déjà des dissyllabes majoritairement iambiques. De plus, Vihman et al. (2006) ont observé que la production des enfants ayant le français comme langue maternelle aurait une plus grande correspondance avec le modèle adulte que celle d'enfants apprenant l'anglais ou le gallois, puisque ces langues comportent des schémas accentuels moins homogènes que le français; c'est-à-dire qu'on trouve des structures à la fois iambiques et trochaïques en gallois et en anglais, mais uniquement des structures iambiques en français.

Ces mouvements globalement de plus grande amplitude, de vitesse réduite et de plus grande durée, pour les enfants par rapport aux adultes, peuvent être envisagés comme des indices d'une production hyperarticulée. En lien avec cette interprétation, quelques études montrent que les enfants ne possèdent pas la capacité de réduire les voyelles inaccentuées, résultant en une production globalement hyperarticulée. Allen et Hawkins (1980) ont étudié des enfants anglophones âgés de un à trois ans et ont montré que les voyelles inaccentuées produites par ce groupe sont plus périphériques, dans l'espace acoustique, que les voyelles inaccentuées produites par les adultes. En conséquence, ces auteurs proposent que le développement de l'accentuation pour les enfants consiste en la maîtrise de la réduction des syllabes inaccentuées. Dans la présente étude, nous évaluons l'hypothèse selon laquelle les enfants francophones n'exploiteront pas la dimension hypo-hyperarticulation autant que les adultes, en condition d'emphase contrastive par rapport à la condition neutre. En conséquence, chez l'enfant, peu de déplacements des articulateurs labiaux et mandibulaires de même que peu de différences dans l'espace F1 vs. F2 seront observées entre les deux conditions prosodiques. Puisque nous relierons l'hypoarticulation au développement de la parole, cette étude permettra de déterminer également si la maîtrise fine du déplacement des articulateurs labiaux et mandibulaires est complétée à l'âge de quatre et de cinq ans².

4. MÉTHODOLOGIE

Afin d'atteindre les objectifs mentionnés à la section précédente, des enregistrements articulatoires et acoustiques ont été menés.

4.1. Sujets et corpus

Cinq adultes âgés entre 25 et 35 ans (trois femmes, deux hommes) ainsi que cinq enfants âgés de quatre et de cinq ans (trois filles, deux garçons) ont été sélectionnés.

²Il ne sera pas discuté ici de la notion de « contrôle » des articulateurs. En effet, bien peu de données nous permettent, à l'heure actuelle, de relier le comportement observé d'un articulateur chez l'enfant à ses origines motrices neuronales et à l'interaction entre les diverses actions du conduit vocal (Fitch 2002).

Les sujets avaient tous le français québécois comme langue maternelle et d'usage. Aucun sujet n'avait éprouvé par le passé de troubles de production ou de perception de la parole. Le test d'évaluation des compétences langagières EVIP P-Body (Échelle de Vocabulaire en Images Peabody; Dunn et al. 1993) a été administré aux cinq locuteurs enfants, qui ont obtenu des résultats conformes à la moyenne des enfants sans trouble développemental du langage. La séquence cible était constituée du disyllabe [baba], intégré dans une phrase porteuse produite dans deux conditions prosodiques : sous emphase contrastive et en condition neutre. Dans la condition neutre, la phrase porteuse était *J'ai vu Baba qui mange un gâteau*, dans laquelle [baba] est en condition neutre, alors que la phrase porteuse en condition d'emphase contrastive était *Non, c'est BABA qui mange un gâteau*, dans laquelle [baba] est en condition d'emphase contrastive. Ces deux phrases faisaient partie d'une histoire racontée par l'expérimentateur. Dans chacune des deux conditions, le locuteur avait pour tâche de décrire l'action réalisée par un des personnages. Dans la condition d'emphase, l'expérimentateur demandait confirmation au locuteur en remplaçant le mot [baba] par un autre mot (par exemple, *C'est Bobby qui mange un gâteau?*). Le locuteur répondait alors par la phrase cible dans laquelle [baba] était produit en condition d'emphase contrastive (*Non, c'est Baba qui mange un gâteau*)³. Entre huit et dix répétitions de chacune des phrases cibles ont été obtenues. Notons que la fonction déictique, signalée par l'emphase contrastive suscitée ici, est bien maîtrisée par les enfants d'âge préscolaire (Hornby et Hass 1970). Dans la condition neutre, il peut être prédit qu'un accent primaire sera réalisé sur la seconde syllabe du mot cible [baba]. Cette structure prosodique diffère de celle de la condition d'emphase contrastive, pour laquelle un accent sera réalisé le plus fréquemment sur la première syllabe du mot ([BAba]), sur la seconde syllabe du mot ([baBA]) ou sur les deux syllabes du mot disyllabique ([BABA]). Pour la présente étude, tous les sujets ont produit un patron d'accentuation iambique (cvCV) sur la séquence [baba] en condition neutre, et un patron d'accentuation trochaïque (CVcv) sur la séquence [baba] en condition d'emphase.

³En raison de la forme du corpus (grandement contrainte par le fait que les locuteurs étaient des enfants d'âge préscolaire), des effets de contexte peuvent être présents. En effet, la première voyelle /a/, dans la condition neutre, était précédée d'une voyelle arrondie /y/ alors que la première voyelle /a/ dans la condition d'emphase était précédée de la voyelle non arrondie /e/. Les deux voyelles /y/ et /e/ diffèrent en ce qui concerne la forme labiale, /y/ étant produite avec les lèvres plus projetées et fermées que /e/. Si aucune consonne n'était produite entre la voyelle /y/ et la première voyelle du mot [baba], un phénomène de coarticulation labiale aurait influencé le degré de protrusion de la voyelle /a/ (la projection des lèvres pour /y/ se poursuivant dans le segment /a/ suivant), diminuant ainsi l'aire aux lèvres qui y est associée. Cependant, /y/ est suivie d'une consonne occlusive bilabiale /b/ qui tend à rétablir le degré de protrusion labiale. En conséquence, il est fort peu probable que les différences de protrusion entre les voyelles /e/ et /y/ puissent être à l'origine des différences observées entre les conditions.

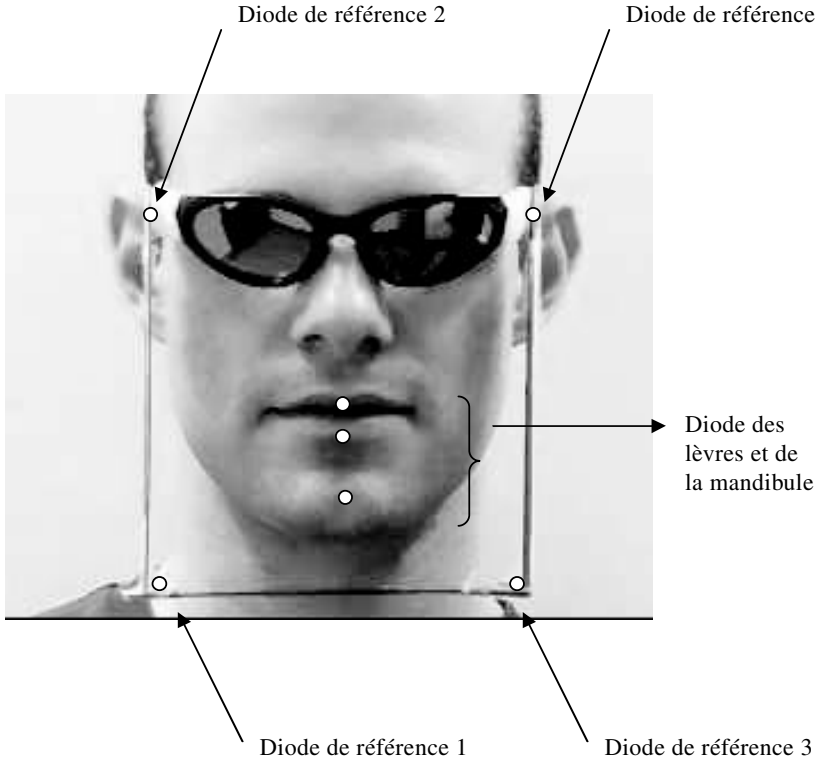


Figure 1: Représentation schématique du dispositif expérimental

4.2. Procédure expérimentale

Les coordonnées des lèvres supérieure et inférieure ainsi que de la mâchoire inférieure ont été enregistrées à l'aide d'un Optotrak. Ce système, qui capte les déplacements tridimensionnels de diodes par caméra infrarouge, est doté d'une précision de mesure de 0,1mm. Trois diodes de 6mm de diamètre ont été posées à l'aide d'adhésif médical sur la ligne centrale du vermillon des lèvres supérieure (Lsup) et inférieure (Linf). Une diode semblable a été posée sur le menton au point correspondant à la base de la gencive (Mand). Afin de corriger les trajectoires en fonction des mouvements de la tête, un cadre de référence rigide, sur lequel étaient posées quatre diodes, a été installé sur des lunettes portées par le locuteur. La figure 1 présente le dispositif expérimental, pour un locuteur adulte. Les coordonnées x , y et z de chacune des sept diodes (quatre diodes de référence et trois diodes posées sur les articulateurs cibles) ont été enregistrées à un taux d'échantillonnage de 175Hz, par le système Optotrak. Le signal audio a été enregistré à l'aide d'un micro de haute qualité à une fréquence de 10000Hz et synchronisé avec les trajectoires articuloaires. Les trajectoires des diodes posées sur les lèvres et la mandibule ont été corrigées en utilisant une technique de soustraction tridimensionnelle préalablement à toute analyse articuloaire.

4.3. Analyses articulatoires et acoustiques

Pour chacune des séquences cibles [baba], les données articulatoires et acoustiques ont été analysées de la façon suivante. La distance interlabiale a été obtenue en soustrayant, dans la dimension verticale (y), la position de la diode posée sur la lèvre inférieure de la position de la diode posée sur la lèvre supérieure ($\text{Inter} = \text{Lsup}_y - \text{Linf}_y$). Les valeurs maximales de distance interlabiale correspondant aux voyelles ouvertes [a] ont été identifiées. Les positions verticales (y) des diodes posées sur les lèvres supérieure et inférieure et sur la mandibule ont été identifiées au même instant T. Ces événements articulatoires sont représentés, à la figure 2, sur une séquence type. À l'aide de ces événements, les paramètres suivants ont été obtenus, pour la séquence [baba] :

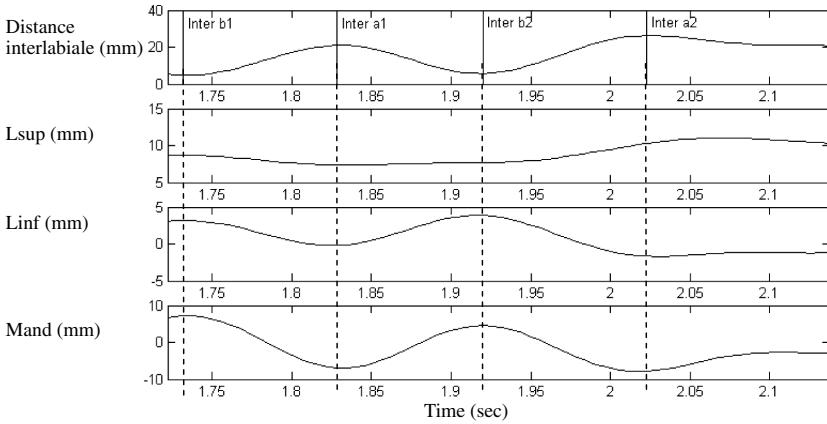
- (i) distance interlabiale minimale pour chacune des consonnes [b] (Inter_{b_1} , Inter_{b_2});
- (ii) distance interlabiale maximale pour chacune des voyelles [a] (Inter_{a_1} , Inter_{a_2});
- (iii) position, dans la dimension y , de la lèvre supérieure, de la lèvre inférieure et de la mandibule aux temps T correspondant aux événements Inter_{b_1} , Inter_{a_1} , Inter_{b_2} et Inter_{a_2} (ces événements seront étiquetés Lsup_{b_1} , Lsup_{a_1} , Lsup_{b_2} , Lsup_{a_2} , Linf_{b_1} , Linf_{a_1} , Linf_{b_2} , Linf_{a_2} , Mand_{b_1} , Mand_{a_1} , Mand_{b_2} , Mand_{a_2}); et
- (iv) déplacement de chacun des trois articulateurs (Lsup, Linf et Mand) lors de la transition de [b] à [a] pour chacune des deux syllabes, calculé par la différence entre la position y associée à la voyelle et la position y associée à la consonne :

$$\begin{aligned} & \text{Lsup}_{a_1} - \text{Lsup}_{b_1}, \text{Lsup}_{a_2} - \text{Lsup}_{b_2} \\ & \text{Linf}_{a_1} - \text{Linf}_{b_1}, \text{Linf}_{a_2} - \text{Linf}_{b_2} \\ & \text{Mand}_{a_1} - \text{Mand}_{b_1}, \text{Mand}_{a_2} - \text{Mand}_{b_2} \end{aligned}$$

De plus, les valeurs formantiques mesurées au centre de la première et de la seconde voyelle [a] du mot cible [baba] ont été mesurées de façon semi-automatique par un algorithme de calcul par *Linear Predictive Coding*, à l'aide du logiciel Praat. Le nombre de pôles variait de 3 à 4. Une fenêtre de type Hamming d'une largeur de 14ms a été utilisée, avec un facteur de préemphasis de 0.98, à partir de 50Hz. Une inspection visuelle a été effectuée pour vérifier l'exactitude des valeurs formantiques fournies par l'algorithme. Ainsi, pour chacune des séquences, les valeurs formantiques ont été superposées à un spectrogramme à bande large. Lorsque la valeur formantique déterminée par l'algorithme ne correspondait pas à la valeur formantique déterminée par inspection visuelle du spectrogramme, l'analyse était exécutée à nouveau, en faisant varier le nombre de pôles. Les valeurs de F0 ont également été extraites au centre de la voyelle par un algorithme d'autocorrélation. Afin de représenter plus fidèlement le mode de traitement par l'oreille humaine de l'échelle des Hertz, les valeurs ont été converties en échelles psycho-acoustiques. Conformément aux études antérieures, les valeurs de F0, en Hertz, ont été transformées en demi-tons (dt) selon la formule suivante :

$$F(dt) = \frac{(12 * \ln(\frac{x}{100}))}{\ln 2}$$

a.



b.

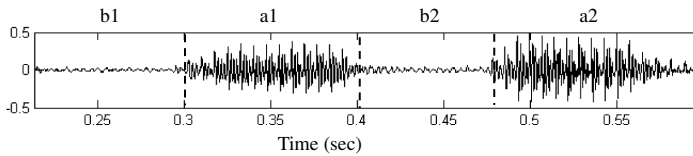


Figure 2: Événements articulatoires (a) et acoustiques (b) étiquetés sur les signaux

où \ln = logarithme népérien et x = fréquence en Hertz.

Les valeurs de F1 et F2 ont été transformées en mels, selon la formule suivante :

$$F(\text{mels}) = 550 * \ln\left(1 + \left(\frac{x}{550}\right)\right)$$

où \ln = logarithme népérien et x = fréquence en Hertz.

Enfin, l'intensité sonore du signal acoustique (amplitude RMS — *root mean square*) de même que la durée ont été mesurées pour chacune des voyelles.

4.4. Analyses statistiques

Des analyses de variance à mesures répétées (ANOVA) ont été effectuées sur chacune des variables dépendantes mentionnées à la section 4.3 (distance interlabiale (Inter), déplacement de chacun des trois articulateurs lors de la transition de [b] à [a], durée des voyelles, premier et second formants, intensité et F0), selon la procédure proposée par Max et Onghena (1999). Les variables indépendantes, qui constituaient les facteurs intra-sujets, étaient les suivantes : *condition prosodique* (emphasis et neutre) et *position dans le mot* (première syllabe ou seconde syllabe). La variable *groupe d'âge* (enfants et adultes) constituait le facteur inter-sujets.

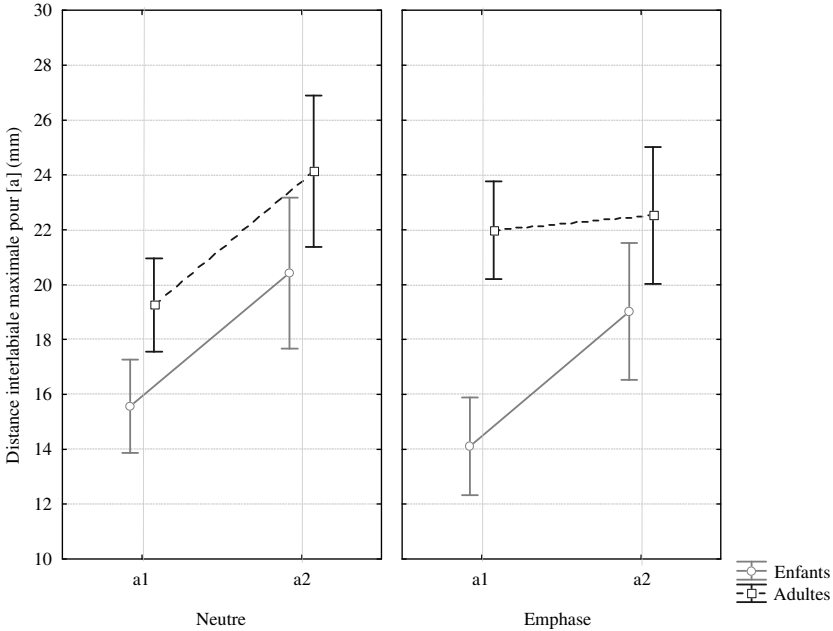


Figure 3: Valeurs moyennes et écarts-types de la distance interlabiale maximale pour les deux voyelles [a] de la séquence [baba] articulées en condition d'emphase contrastive et en condition neutre

5. RÉSULTATS

Les résultats sont présentés pour chacune des variables dépendantes étudiées. Les résultats détaillés des analyses statistiques ne sont donnés que lorsque l'une ou l'autre des variables indépendantes a un effet significatif sur la variable dépendante à l'étude.

5.1. Distance interlabiale maximale pour [a]

Les valeurs moyennes de la distance interlabiale maximale réalisée pour les voyelles [a] de la séquence [baba] sont représentées par le graphique de la figure 3.

Les résultats des analyses statistiques ont révélé que pour les deux groupes de sujets, la distance interlabiale associée à la voyelle de la seconde syllabe est plus grande que la distance interlabiale correspondant à la voyelle de la première syllabe ($F(1, 8) = 44, 43; p < 0, 05$). Un effet significatif d'interaction entre les variables *groupe*, *condition* et *position* est ressorti. Pour les adultes seulement, la distance interlabiale maximale pour [a] en première position est plus grande en condition d'emphase qu'en condition neutre ($F(1, 8) = 5, 55; p < 0, 05$). Aucune différence n'est observée pour le groupe des enfants. Ces résultats démontrent que la production d'un accent d'emphase chez les enfants étudiés n'implique pas une augmentation de la distance interlabiale par rapport à la condition neutre. Par contre, la voyelle [a]

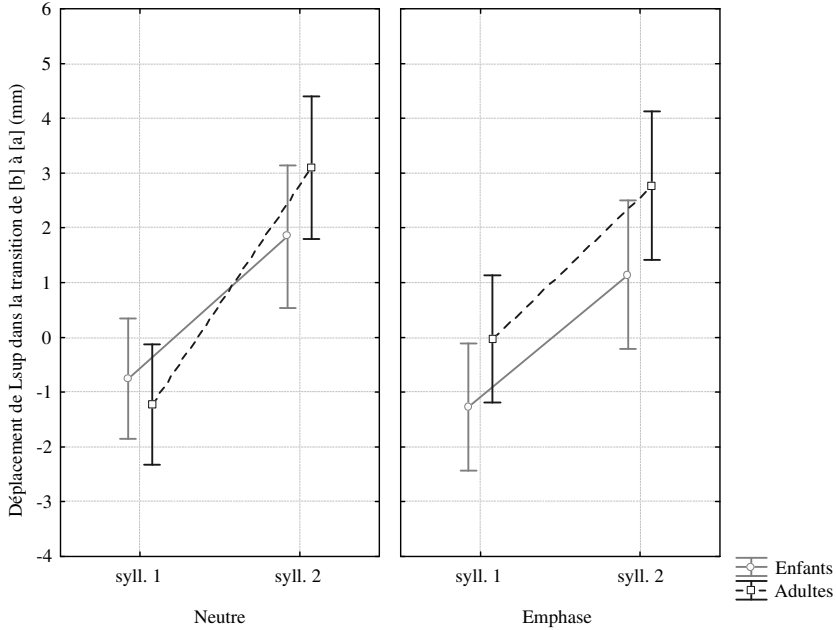


Figure 4: Valeurs moyennes et écarts-types du déplacement de la lèvres supérieure lors de la transition de [b] à [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

sous accent d'emphase est produite par une augmentation de la distance interlabiale chez les adultes.

5.2. Déplacement de la lèvre supérieure

La figure 4 représente les valeurs de déplacement de la lèvre supérieure lors de la transition de [b] à [a], pour chacune des deux syllabes de la séquence [baba].

Les analyses statistiques révèlent que pour les deux groupes de sujets, dans les deux conditions prosodiques, le déplacement de la lèvre supérieure produit pour la seconde syllabe est significativement plus grand que celui produit pour la première syllabe ($F(1, 8) = 59, 53; p < 0, 05$). Cependant, un effet significatif de l'interaction entre les variables *groupe*, *condition* et *position* est trouvé. Pour les adultes, le déplacement de la lèvre supérieure produit pour la syllabe [ba] en première position est significativement plus grand en condition d'emphase qu'en condition neutre ($F(1, 8) = 6, 69; p < 0, 05$). En revanche, sur cette même syllabe, aucune différence significative n'est produite par les enfants entre les déplacements en condition d'emphase par rapport à la condition neutre. Sur cette syllabe, les déplacements (dans les deux conditions prosodiques) produits par les enfants sont comparables à ceux produits par les adultes, ce qui tend à confirmer que les enfants produisent des mouvements globalement hyperarticulés, compte tenu de la petite taille de leur cavité. En seconde position, les déplacements de la lèvre supérieure sont plus grands que ceux

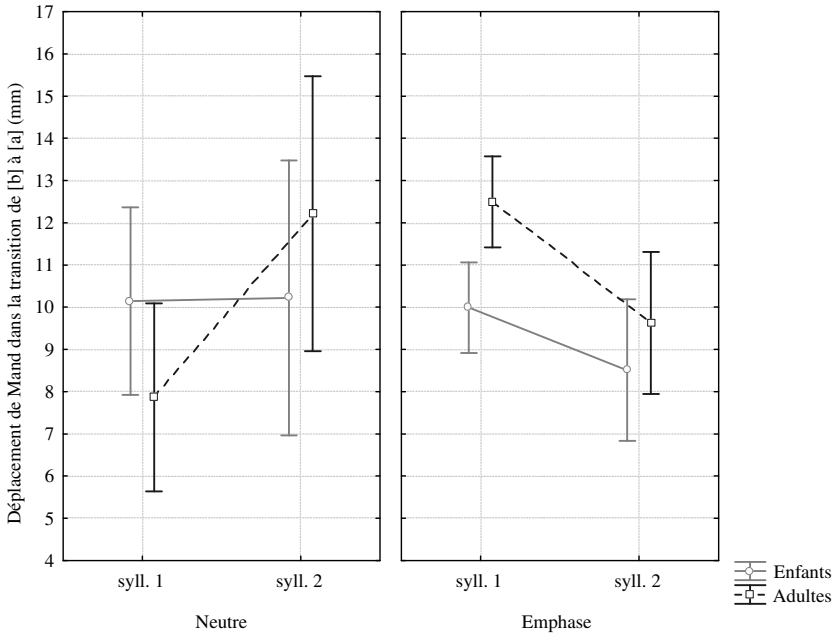


Figure 5: Valeurs moyennes et écarts-types du déplacement de la mandibule lors de la transition de [b] à [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

produits en première position, et ce pour les deux groupes de sujets et dans les deux conditions. Ce résultat montre que l'effet de frontière droite du domaine du syntagme intonatif est toujours présent dans les deux conditions prosodiques.

5.3. Déplacement de la mandibule

Les valeurs moyennes et écarts-types du déplacement de la mandibule lors de la transition de [b] à [a] pour les deux syllabes de la séquence produite dans les deux conditions prosodiques sont représentées à la figure 5.

Les résultats des ANOVA ont montré que l'effet d'interaction entre les variables *condition prosodique* et *position* était significatif ($F(1, 8) = 15,69; p < 0,05$). Globalement (sujets confondus), en première position, les valeurs de déplacement sont significativement plus grandes pour la condition d'emphase contrastive que pour la condition neutre ($F(1, 8) = 9,03; p < 0,05$). Les patrons de variation de cette variable diffèrent cependant en fonction du groupe de sujets considéré (effet significatif de l'interaction entre *condition*, *position* et *groupe* ($F(1, 8) = 6,55; p < 0,05$)). Pour les sujets adultes, en syllabe initiale, les déplacements de [b] à [a] sont significativement plus grands en condition d'emphase qu'en condition neutre ($F(1, 8) = 19,34; p < 0,05$). Pour le groupe des enfants, les déplacements ne varient pas de façon significative entre les conditions emphase et neutre. En seconde syllabe, les

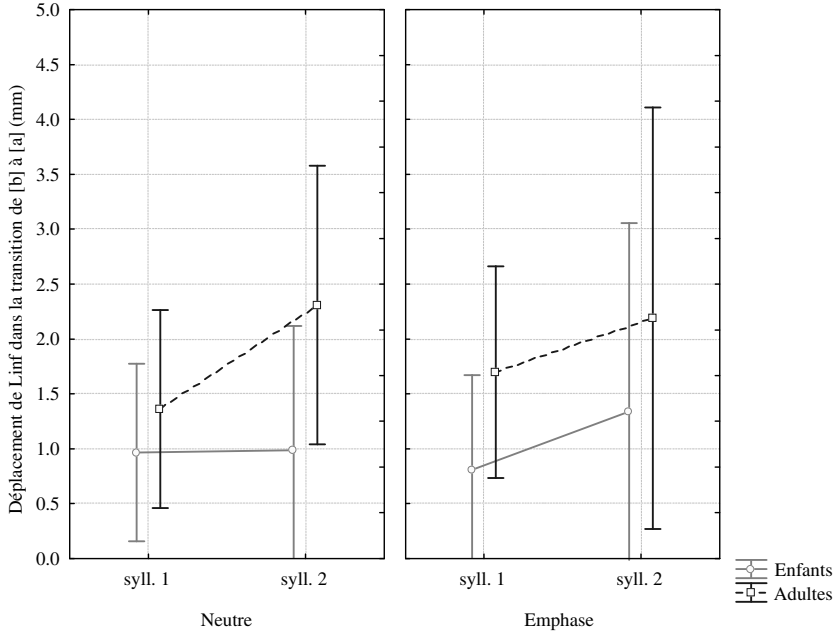


Figure 6: Valeurs moyennes et écarts-types du déplacement de la lèvre inférieure (sans la mandibule) lors de la transition de [b] à [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

déplacements ne varient pas de façon significative en fonction des facteurs *position* et *groupe*. Conformément aux résultats de la section précédente, l'absence de différences entre les valeurs de déplacements des deux groupes de sujets peut être interprétée comme un indice que la parole des enfants est de façon globale hyperarticulée, compte tenu de la petite taille de leur conduit vocal.

5.4. Déplacement de la lèvre inférieure

La figure 6 représente les valeurs de déplacement de la lèvre inférieure lors de la transition de [b] à [a], pour les deux groupes de sujets, dans les deux conditions prosodiques et pour les deux positions syllabiques.

Notons que la trajectoire de la diode posée sur la mandibule a été soustraite de la trajectoire de la diode posée sur la lèvre inférieure afin de calculer le mouvement propre de la lèvre inférieure. Les analyses statistiques n'ont pas révélé d'effet significatif de la *condition*, de la *position* ou du *groupe* de sujets sur les valeurs de déplacement. Il est à noter par ailleurs que la valeur élevée des écarts-types témoigne de l'importante variabilité des données.

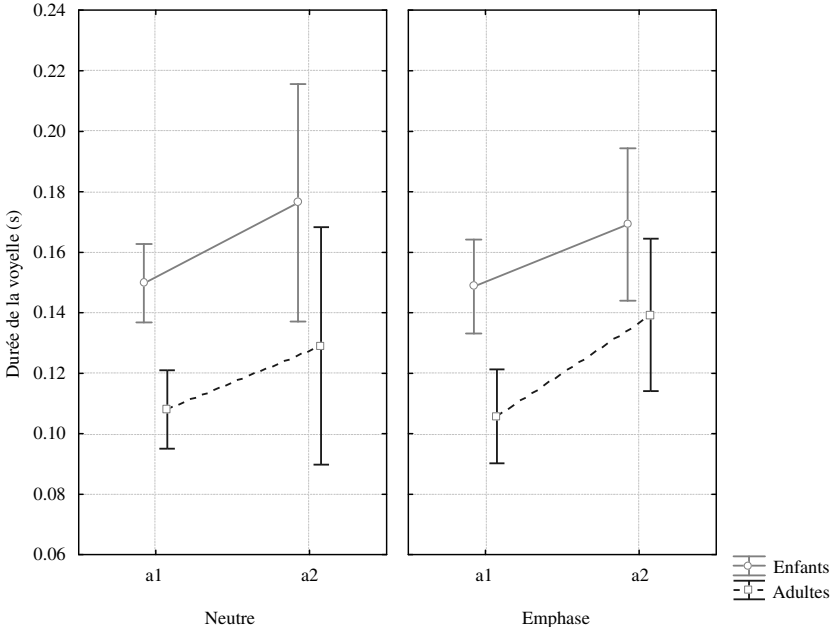


Figure 7: Valeurs moyennes et écarts-types de la durée de la voyelle [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

5.5. Durée vocalique

La durée de la voyelle [a], telle que déterminée sur le signal acoustique, est représentée en fonction des variables indépendantes à la figure 7. Les analyses statistiques révèlent un effet significatif de la variable *groupe* ($F(1, 8) = 10,97; p < 0,05$), les enfants produisant des voyelles significativement plus longues que celles des adultes. La variable *position* a également un effet significatif sur la valeur de la durée ($F(1, 8) = 11,57; p < 0,05$). Comme le suggère la figure 7, les voyelles produites en deuxième position du mot [baba] (a2) sont significativement plus longues que celles produites en première position (a1). Ce résultat suggère que les corrélats temporels de l'accent primaire sont produits dans les deux conditions prosodiques (neutre et emphase contrastive). Par ailleurs, aucun effet d'interaction n'est relevé entre les trois variables, ce qui confirme que ce patron de variation est similaire pour les deux groupes de sujets et dans les deux conditions prosodiques.

5.6. Valeurs formantiques

Les figures 8 et 9 regroupent les valeurs des deux premiers formants des voyelles [a] pour les deux groupes de sujets. Comme le suggèrent les graphiques, les valeurs formantiques produites par les enfants sont significativement plus élevées que celles produites par les adultes, en F1 ($F(1, 8) = 11,88; p < 0,05$) et en F2

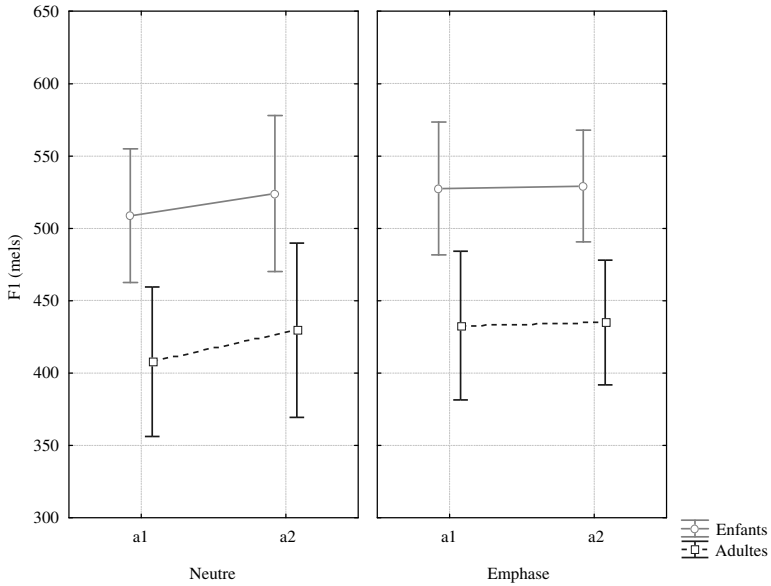


Figure 8: Valeurs moyennes et écarts-types de la valeur du premier formant (F1), en mels, de la voyelle [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

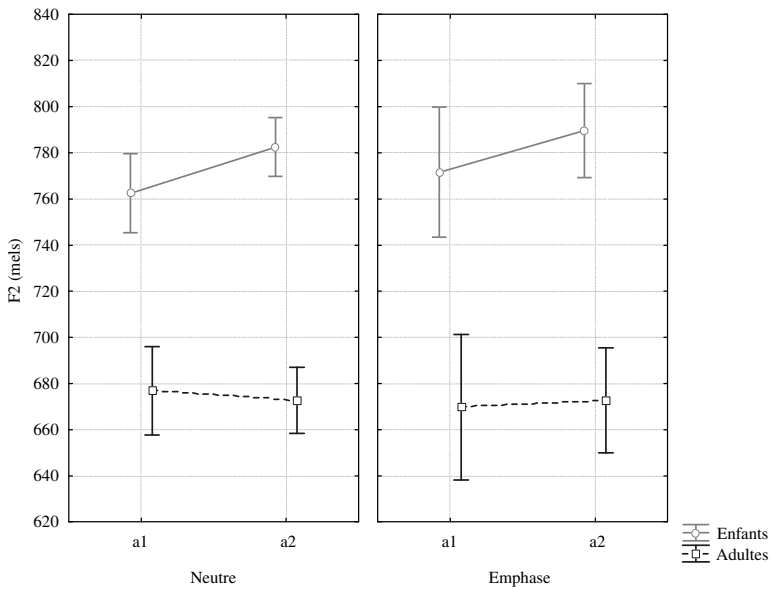


Figure 9: Valeurs moyennes et écarts-types de la valeur du deuxième formant (F2), en mels, de la voyelle [a] pour chacune des deux syllabes de la séquence [baba], en condition d'emphase contrastive et en condition neutre

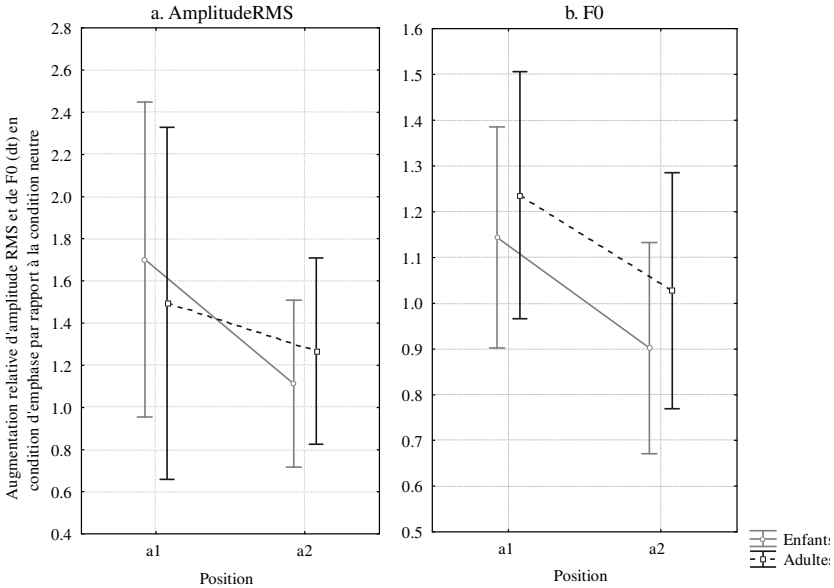


Figure 10: Augmentations relatives (moyennes et écarts-types) de la valeur d'amplitude (gauche) et de F0 (droite) de la voyelle [a] pour chacune des deux syllabes de la séquence [baba], entre la condition d'emphase contrastive et la condition neutre

($F(1, 8) = 103, 01; p < 0, 05$). Ce résultat est prévisible compte tenu de la taille réduite de la cavité pour les enfants par rapport aux adultes. Par ailleurs, aucun effet significatif de la condition prosodique ou de la position des voyelles dans le mot cible [baba] n'a été noté. Il apparaît donc que les effets articulatoires de l'emphase contrastive, tels que démontrés dans les sections précédentes, ne donnent pas systématiquement lieu à des conséquences sur la valeur des formants. Il va sans dire que d'autres articulateurs, comme la langue, sont en jeu dans la production de la voyelle [a], articulateurs qui ne font pas l'objet d'étude ici.

5.7. Valeurs de F0 et d'amplitude

Les valeurs relatives d'amplitude et de F0 (en demi-tons) sont présentées à la figure 10. Pour chacune des positions syllabiques du mot [baba], les valeurs moyennes mesurées pour la voyelle [a] en condition d'emphase ont été divisées par les valeurs moyennes pour cette voyelle en condition neutre. Les données de la figure 10 représentent donc l'augmentation relative d'amplitude et de F0 pour chaque position en condition d'emphase par rapport à la condition neutre. Une valeur supérieure à 1 signifie que la valeur en condition d'emphase est plus élevée que la valeur en condition neutre.

En ce qui concerne l'amplitude (figure 10a), les analyses statistiques révèlent que les valeurs en première position (a1) sont significativement plus élevées que

les valeurs en seconde position (a2) ($F(1, 8) = 6,07; p < 0,05$). Aucun effet du groupe n'est relevé, ce qui signifie que les enfants et les adultes augmentent de façon comparable l'amplitude de la première voyelle en condition d'emphase. Un patron similaire est observable pour les valeurs de F0 (figure 10b). La première syllabe est produite avec une augmentation de F0 en condition d'emphase par rapport à la condition neutre ($F(1, 8) = 11,34; p < 0,05$), et ce, pour les deux groupes de sujets. Il apparaît donc que les deux groupes de sujets (enfants et adultes) signalent l'accent d'emphase par une augmentation similaire d'amplitude et de F0 sur la première syllabe du mot [baba]. Pour les deux variables (F0 et amplitude), des écarts-types élevés sont toutefois observables.

6. DISCUSSION

Les résultats présentés jusqu'à maintenant laissent entrevoir l'existence à la fois de similitudes et de différences dans la réalisation phonétique des représentations prosodiques chez les enfants de quatre et de cinq ans, par rapport aux adultes. En effet, pour les deux groupes de sujets, les paramètres d'amplitude et de F0 ont été modifiés par l'accent d'emphase. L'accent d'emphase, qui affecte la première syllabe du mot [baba], a été produit avec une intensité accrue et une montée de F0, conformément aux études antérieures (Jun et Fougeron 2002; Ménard et al. 2006). Ces résultats sont prévisibles, Vihman (1996) ayant démontré que les articulateurs glottiques sont maîtrisés très tôt comparativement aux articulateurs supraglottiques. Cependant, bien que la voyelle [a] en seconde position soit significativement plus longue que sa contrepartie initiale, aucun allongement significatif n'est produit sur la première syllabe en condition d'emphase, par rapport à la condition neutre. Ce patron démontre bien la robustesse de l'accent primaire sur la dernière syllabe du mot, dont les corrélats temporels sont réalisés dans les deux conditions prosodiques.

Néanmoins, nous avons montré que les effets articulatoires de l'accent d'emphase variaient selon le groupe de sujets considéré. D'abord, la voyelle [a] en première position est associée à une distance interlabiale significativement plus grande en condition d'emphase par rapport à la condition neutre pour le groupe des adultes. Aucune différence relative à la condition prosodique n'est observée chez les enfants. Comme le dénotent les figures 4 à 6, cette variation de distance interlabiale est reliée à une augmentation significative du déplacement de la lèvre supérieure et de la mandibule lors de la transition de [b] à [a] chez les adultes. Aucune différence significative n'a été observée pour le groupe des enfants, en ce qui a trait à la production de la voyelle dans cette position. En somme, dans la présente expérience, les enfants ont réalisé l'accent d'emphase par une augmentation d'intensité sonore et de fréquence fondamentale, sans toutefois modifier les trajectoires articulatoires des lèvres et de la mandibule. Le fait que les déplacements labiaux et mandibulaires produits par les jeunes enfants ont été comparables à ceux produits par les adultes, compte tenu des tailles très différentes de cavités vocales, révèle que les enfants produisent des mouvements globalement hyperarticulés, par rapport aux adultes (Ménard et al. 2006). Les durées vocaliques produites par les enfants sont également plus grandes que chez les adultes, dans les deux positions et les deux conditions

prosodiques. Ces résultats sont conformes aux études antérieures, qui soulignent le fait que les enfants d'âge préscolaire, bien qu'ils soient en mesure de réaliser des syllabes perçues comme accentuées par les auditeurs, réalisent les accents de façon différente, par rapport aux adultes. La maîtrise du complexe labio-mandibulaire ne serait complète que vers l'âge de six ans, comme le montre une étude de la variabilité intra-individuelle chez des locuteurs âgés de quatre ans et de six ans et des adultes (Green et al. 2000; Green et al. 2002). L'imaturité des habiletés articulatoires se manifesterait, ici, par une difficulté à réduire les voyelles en condition neutre, par rapport à la condition d'emphase. Il pourrait cependant être proposé que ces différences ne sont que le reflet d'un choix de stratégies plutôt que d'une différence de production reliée aux capacités articulatoires (Esling et al. 2004). Puisque la maîtrise des articulateurs supraglottiques n'est pas acquise par les locuteurs enfants, ceux-ci opéreraient pour la manipulation des dimensions de F0 et d'intensité, plutôt que la manipulation des dimensions labio-mandibulaires. Cette hypothèse serait par ailleurs difficilement conciliable avec la nature même de l'emphase contrastive dans la tâche décrite ici : les locuteurs doivent faire ressortir, par une tâche linguistique, une syllabe par rapport à son contexte environnant. Le but recherché est clairement un effet perceptif (augmentation du contraste perçu sur une syllabe), et dans un souci d'efficacité, le locuteur exploite tous les moyens à sa disposition pour accomplir cette tâche. Ce paradigme aurait donc dû forcer les locuteurs enfants à hyperarticuler les voyelles s'ils en avaient eu la possibilité.

Contrairement aux résultats de Goffman (1999), les accents finaux produits en condition neutre (donc en structure iambique) étaient bien différenciés articulatoirement par rapport à la syllabe inaccentuée précédente. Bien que nos sujets enfants soient plus âgés que ceux de Vihman et al. (1998) et de Vihman et al. (2006), nos résultats confirment les résultats de ces études qui réfutent l'hypothèse d'une acquisition précoce des patrons rythmiques trochaïques par rapport aux patrons rythmiques iambiques. Conformément à Vihman et al. (2006), par ailleurs, la production de la structure rythmique de la langue maternelle ne serait possible que tardivement (après quatre ans, selon nos données).

Le but de la présente étude était d'examiner les effets locaux de l'accent d'emphase sur la variation des patrons articulatoires labiaux-mandibulaires et des paramètres acoustiques. La langue, articulateur dont la maîtrise est atteinte plus tardivement que celle des lèvres et de la mandibule, est un organe crucial dans l'atteinte des cibles acoustiques. Des travaux sont actuellement en cours afin d'étudier, à l'aide d'un échographe, les effets des représentations prosodiques sur le déplacement de la langue, des lèvres et de la mâchoire. Le travail présenté ici s'intègre dans un programme de recherche plus vaste visant à étudier l'effet des contraintes périphériques anatomiques et articulatoires sur l'émergence de la parole chez l'enfant (Ménard et Boë 2004; Ménard et al. 2006). Il apparaît crucial de tenir compte de l'impact de ce type de contraintes dans la mise en forme des unités phonologiques produites par l'enfant.

RÉFÉRENCES

- Allen, G.D. et S. Hawkins. 1980. Phonological rhythm : Definition and development. In *Child Phonology*, vol. 1 : *Production*, sous la dir. de Grace H. Yeni-Komshian, James F. Kavanagh et Charles A. Ferguson, 227–256. New York : Academic Press.
- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht : Foris.
- Beckman, Mary E., Jan Edwards et Janet Fletcher. 1992. Prosodic structure and tempo in a sonority model of articulatory dynamics. In *Papers in Laboratory Phonology II : Segment, Gesture, Prosody*, sous la dir. de Gerard J. Docherty et D. Robert Ladd, 68–86. Cambridge : Cambridge University Press.
- Cho, Taehong. 2002. *The effects of prosody on articulation in English*. New York : Routledge.
- Cho, Taehong. 2005. Prosodic strengthening and featural enhancement : Evidence from acoustic and articulatory realizations of /a, i/ in English. *Journal of the Acoustical Society of America* 117:3867–3878.
- Connaghan, Kathryn P., Christopher A. Moore, K.J. Reilly, K.B. Almand et R.W. Steeve. 2001. Acoustic and physiologic correlates of stress production across systems. Affiche présentée à l'American Speech-Language-Hearing Association Convention, Nouvelle-Orléans, LA.
- De Jong, Kenneth J. 1995. The supraglottal articulation of prominence in English : Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97:491–504.
- Di Cristo, Albert. 1998. Intonation in French. In *Intonation systems : A survey of twenty languages*, sous la dir. de Daniel Hirst et Albert Di Cristo, 195–218. Cambridge : Cambridge University Press.
- Dunn, Lloyd M., Claudia M. Thériault-Whalen et Leota M. Dunn. 1993. Échelle de vocabulaire en images Peabody (ÉVIP). Adaptation française du Peabody Picture Vocabulary Test, Revised. Toronto : Psycan.
- Erickson, Donna. 1998. Effects of contrastive emphasis on jaw opening. *Phonetica* 55:147–169.
- Erickson, Donna. 2002. Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59:134–149.
- Esling, John H. 2005. There are no back vowels : The laryngeal articulator model. *Revue canadienne de linguistique* 50:13–44.
- Esling, John H., Allison Benner, Lisa Bettany et Chakir Zeroual. 2004. Le contrôle articulatoire phonétique dans le prébabillage. In *Actes des XXVes Journées d'Étude sur la Parole*, sous la dir. de Bernard Bel et Isabelle Marlien, 205–208. Fez, Maroc : Association Francophone de la Communication Parlée.
- Fitch, W. Tecumseh. 2002. Comparative vocal production and the evolution of speech : Reinterpreting the descent of the larynx. In *The transition to language*, sous la dir. de Alison Wray, 21–45. Oxford : Oxford University Press.
- Goffmann, Lisa. 1999. Prosodic influences on speech production in children with specific language impairments and speech deficits. *Journal of Speech, Language, and Hearing Research* 42:1499–1517.
- Goffman, Lisa. 2004. Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, and Hearing Research* 47:1088–1102.
- Goffman, Lisa et Caren Malin. 1999. Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research* 42:1003–1015.
- Green, Jordan R., Christopher A. Moore, Masahiko Higashikawa et Roger W. Steeve. 2000. The physiologic development of speech motor control : Lip and jaw coordination. *Journal*

- of *Speech, Language, and Hearing Research* 43:239–255.
- Green, Jordan R., Christopher A. Moore et Kevin J. Reilly. 2002. The sequential development of lip and jaw control for speech. *Journal of Speech, Language, and Hearing Research* 45:66–79.
- Hornby, Peter A. et Wibur A. Hass 1970. Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* 13:395–399.
- Jun, Sun-Ah et Cécile Fougeron. 2000. A phonological model of French intonation. In *Intonation : Analysis, modelling and technology*, sous la dir. de Antonis Botinis, 209–242. Dordrecht : Kluwer.
- Jun, Sun-Ah et Cécile Fougeron. 2002. Realizations of accentual phrase in French intonation. *Probus* 14:147–172.
- Kehoe, Margaret, Carol Stoel-Gammon et Eugene H. Buder. 1995. Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research* 38:338–350.
- Konopczynski, Gabrielle. 1986. Du prélangage au langage : acquisition de la structuration prosodique. Thèse de doctorat, Université de Strasbourg II.
- Lacheret-Dujour, Anne et Frédéric Beaugendre. 1999. *La prosodie du français*. Paris : CNRS Langage.
- Lævenbruck, Héléne. 1999. An investigation of articulatory correlates of the accentual phrase in French. In *Proceedings of the XIVth International Congress of Phonetic Sciences*, sous la dir. de John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville et Ashlee C. Bailey, 667–670. University of California, Berkeley.
- MacNeilage, Peter F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21:499–511.
- Max, Ludo et Patrick Onghena. 1999. Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research* 42:261–270.
- Ménard, Lucie et Louis-Jean Boë. 2004. L'émergence du système phonologique chez l'enfant : l'apport de la modélisation articuloire. *Revue canadienne de linguistique* 49:155–174.
- Ménard, Lucie, Héléne Lævenbruck et Christophe Savariaux. 2006. Articulatory and acoustic correlates of contrastive focus in French children and adults. In *Speech production : Models, phonetic processes and techniques*, sous la dir. de Johnathan Harrington et Marija Tabain, 227–251. New York : Psychology Press.
- Ruark, Jacki L. et Christopher A. Moore. 1997. Coordination of lip muscle activity by 2-year-old children during speech and nonspeech tasks. *Journal of Speech, Language, and Hearing Research* 40:1373–1385.
- Smith, Anne et Lisa Goffman. 1998. Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language, and Hearing Research* 41:18–30.
- Touati, Paul. 1987. Structures prosodiques du suédois et du français : profils temporels et configurations tonales. *Travaux de l'institut de Linguistique de Lund* 21. Lund, Suède : Lund University Press.
- Vihman, Marilyn May. 1996. *Phonological development : The origins of language in the child*. Cambridge, MA : Blackwell.
- Vihman, Marilyn May, Satsuki Nakai et Rory DePaolis. 2006. Getting the rhythm right : A cross-linguistic study of segmental duration in babbling and first words. *Laboratory phonology* 8, sous la dir. de Louis Goldstein, D.H. Whalen et Catherine T. Best, 341–366. New York : Mouton de Gruyter.
- Vihman, Marilyn May, Rory A. DePaolis et Barbara L. Davis. 1998. Is there a "trochaic bias" in early word learning? Evidence from infant production in English and French. *Child Development* 69:933–949.