

Développement du contrôle des cadres de la parole : une étude longitudinale du contrôle oral/nasal et de la coordination oro-laryngée.

Lalevée, C. & Vilain, A.

Institut de la Communication Parlée(ICP) - CNRS UMR 5009
Université Stendhal, BP 25, Grenoble cedex 9, France
e-mail : lalavec@icp.inpg.fr, avilain@icp.inpg.fr

ABSTRACT

Canonical babbling can be related to the sudden emergence of control over the carrier articulator, i.e. the mandible, in the absence of any voluntary control over the other speech articulators such as the velum, the lips and the tongue. Although it is considered as a major step in speech development, this proto-syllabic frame will not be adapted to the needs of human adult communication until the baby is capable of controlling globally its vocal tract from the glottis to the lips, in order to achieve an efficient acoustic intra-syllabic contrast. Two important elements in this global control are that of the velum, which enables to have a fully oral vocal tract, therefore to produce salient consonant-vowel sequences, and the oro-laryngeal coordination. We have studied the evolution of these controls in speech development from an audio-visual digital video corpus of 2 French children.

1. INTRODUCTION

Le babillage canonique, en tant qu' étape majeure du développement de la parole, se trouve au centre d'un grand nombre d'études depuis quelques années. Il est défini comme l'émergence de proto-syllabes, due au contrôle de séries de mouvements d'ouverture et de fermeture du conduit vocal : le bébé est alors capable de coordonner les gestes glottiques et supra-glottiques à l'intérieur d'une unité rythmique, ce qui constitue la structure de base de la communication humaine. Dans sa théorie "Frames, then Content" de l'évolution de la parole, MacNeilage [15] a proposé de considérer ces cadres proto-syllabiques comme résultant de simples oscillations mandibulaires, sans aucun contrôle des autres articulateurs portés (principe de la "dominance des cadres"). Par la suite, de nombreuses études se sont attachées à tester le poids réel des oscillations mandibulaires, et à montrer dans quelle mesure ce simple mouvement peut expliquer les productions du babillage. Plusieurs études [5], [18], [24] ont évalué la corrélation entre le lieu d'articulation de la consonne et le lieu d'articulation de la voyelle à l'intérieur d'une même syllabe du babillage : une forte corrélation indique une importante interdépendance entre les deux parties de la séquence protosyllabique, que l'on peut attribuer au fait que ces deux parties soient le produit d'un seul mouvement de la porteuse mandibulaire. Une autre étude [25] a testé la possibilité que les oscillations mandibulaires puissent bien être la cause unique des productions du babillage, en les simulant à l'aide de modèles articulatoires anthropomorphiques, pour lesquels tous les degrés de liberté, excepté celui de la mandibule, étaient gelés. Les résultats obtenus confortent la théorie.

Cependant, si l'on peut donc bien considérer le contrôle des oscillations mandibulaires comme la structure de base de la parole, il nous semble important d'ajouter à la dominance des cadres, le contrôle de la closance, qui permettra à l'enfant, à travers, le contrôle du velum d'une part, et le contrôle de la coordination oro-laryngée d'autre part, d'acquiescer un contrôle global du conduit vocal, de la glotte aux lèvres, afin d'obtenir un contraste acoustique efficace à l'intérieur de la syllabe. Le contrôle du velum est important pour obtenir un conduit vocal oral qui permet un bon découplage entre consonnes faiblement sonores et voyelles très sonores, alors que le contrôle de la coordination oro-laryngée permet d'obtenir la distinction voisé / non-voisé, pertinente dans la plupart des langues du monde.

2. LE CONTROLE DE LA CLOSANCE

2.1. Un corpus audio-visuel

Pour mener notre étude sur le développement du contrôle du conduit vocal, nous avons travaillé à partir d'un corpus audio-visuel. Ce corpus (VHS + DAT, puis DV) a été recueilli par S. Brosda [2] et comprend les productions de 6 enfants français âgés de 6 à 15 mois. Les enfants ont été enregistrés dans leur environnement familial à raison de 45 minutes tous les quinze jours à partir du moment où les parents ont signalé le début du babillage et pendant les 9 mois suivants. L'étude présentée ici porte sur les données de deux enfants entre 6 et 12 mois, du babillage aux premiers mots. Les vocalisations retenues pour la transcription sont celles qui sont considérées comme de la production de parole et qui correspondent aux critères du babillage [8] (pas de rires, de toux ou de soupirs). Le corpus a été transcrit par deux experts [12], à l'aide du système AMSTIVOC (Amsterdam System for Transcription of Infant VOCalizations [9], [1]) qui décrit de façon détaillée les lieux et modes articulatoires et phonatoires, sous le format CLAN (Computerized Language ANALysis), un programme de la base de données CHILDES (un système d'échange de données de productions d'enfants [16,17]) qui permet de multiples analyses. La transcription des fichiers audio était complétée par la vidéo et basée sur une analyse acoustique à l'aide du logiciel de traitement de la parole PRAAT.

2.2. Le contrôle oral/nasal : une étape majeure dans le développement de la parole.

Les consonnes nasales représentent 14.6% des consonnes des langues du monde (cf. la base de données UPSID 451, [22]). Cette faible proportion peut s'expliquer par la mauvaise saillance perceptive des nasales dans la syllabe. En effet, dans la plupart des langues du monde, la syllabe est caractérisée par un patron d'alternance entre faible sonorité et forte sonorité. Une séquence constituée d'une consonne nasale suivie d'une voyelle présente un faible contraste acoustique car la consonne nasale comme la voyelle possède une large amplitude ainsi qu'une structure formantique. D'autre part, les lieux d'articulation des nasales sont difficilement reconnaissables (10,4% des lieux reconnus [4]) et ceci peut expliquer le fait que les langues exploitent peu de lieux d'articulation différents pour les nasales.

Les sons oraux sont produits avec le velum en position relevée contre la paroi pharyngale, ce qui implique un mouvement actif de contraction de ce muscle. On peut donc se poser la question suivante : la position abaissée du velum est-elle la position par défaut, celle qui caractérise une absence de contrôle dans les premiers temps du développement ? Un certain nombre d'études [14, 21] ont effectivement montré que les vocalisations du bébé contenaient une large proportion de nasales. On peut alors se demander à partir de quel moment et de quelle manière le bébé va acquérir ce mouvement particulier du velum qui permet la fermeture du canal vélo-pharyngé. Mais il faut avant tout préciser la nature de la nasalité et voir si ce contrôle pourra s'étudier de la même manière pour les consonnes et les voyelles.

Différents contrôles pour les consonnes nasales et les voyelles ?

Les consonnes et les voyelles nasales ont des caractéristiques acoustiques et articulatoires bien distinctes. L'ouverture du conduit nasal opérée par le mouvement d'abaissement du velum, a des conséquences très différentes au niveau acoustique, selon que le conduit vocal est ouvert ou fermé. En effet, si le conduit vocal est fermé comme pour les occlusives, les fosses nasales vont agir comme des résonateurs et entraîner l'apparition de formants, ce qui signifie qu'un changement radical s'effectue dans la structure acoustique de la consonne. Par conséquent, la différence entre les deux positions du velum dans une consonne engendre des signes évidents et bien distinguables sur le signal de parole. Par contre, abaisser le velum alors que le conduit vocal est ouvert, i.e. passer d'une voyelle orale à une voyelle nasale, ne modifie pas de manière radicale la nature acoustique du son. L'élargissement du passage vélopharyngé se traduira plutôt par de minces effets tels que l'élargissement de la largeur de bande du premier formant, l'abaissement de l'amplitude du premier formant, ainsi que l'affaiblissement de la fréquence du second formant. Actuellement, la détection de corrélats acoustiques reliés à la nasalité dans les voyelles représente un obstacle pour les recherches acoustiques [7, 8]. La détection de consonnes nasales plutôt que celle des voyelles nasales semble donc plus aisée, surtout lorsqu'il s'agit de vocalisations de bébés.

Un second argument plus fondamental en faveur des consonnes plutôt que des voyelles pour étudier la nasalité a

été avancé par Rossato et al. [20]. Ils ont mené une étude sur les mouvements du velum, plus particulièrement sur la distribution des positions verticales de celui-ci lors de la production de voyelles orales et nasales et de consonnes orales et nasales, chez un locuteur français à l'aide de mesures obtenues avec un articulographe électromagnétique. Il ressort de cette étude que la différence entre une consonne orale et une consonne nasale peut être obtenue par un mouvement du velum de 2 mm, alors que le passage d'une voyelle nasale à une voyelle orale implique une plus large ouverture du conduit nasal et un plus large mouvement d'abaissement du velum (la différence significative entre la position abaissée et la position relevée du velum est de 7 mm). Il apparaît, en outre, qu'une voyelle orale peut être produite avec une faible ouverture du conduit nasal, essentiellement pour les voyelles basses. En d'autres termes, une petite ouverture du passage vélopharyngé n'entraînera pas de changements acoustiques radicaux pour une voyelle orale alors que dans le cas d'une consonne, la structure acoustique changerait radicalement. Une séquence [mama] pourrait donc être produite avec une même position abaissée du velum. En résumé, il semble que l'adulte utilise trois positions différentes : une fermeture complète pour les consonnes et les voyelles orales, une légère ouverture pour les consonnes nasales et les voyelles orales basses et une grande ouverture pour les voyelles nasales.

On peut donc émettre l'hypothèse que dans les premières étapes du développement de la parole, la position par défaut du velum serait celle d'une légère ouverture du passage vélopharyngé telle que pour une consonne nasale, alors que la position basse du velum pour les voyelles nasales impliquerait un contrôle complexe de celui-ci, qui se développerait tardivement dans l'ontogénèse.

Nous avons donc étudié l'évolution de la proportion des nasales dans le babillage afin de déterminer à partir de quel âge le bébé atteignait les normes adultes, ainsi que l'évolution de la nasalité à partir des proportions des consonnes nasales chez deux bébés de 6 à 12 mois.

Indices d'un contrôle actif

Pour conforter le principe de la dominance des cadres, Redford et al. [19] ont mené une étude sur les contraintes de production des consonnes en finale de syllabe dans le babillage, du début du babillage jusqu'à l'âge de douze mois. Ils se sont intéressés, en particulier, aux changements de modes oral à nasal à l'intérieur d'une syllabe. Les changements observés s'effectuent majoritairement dans le sens oral à nasal. Quant aux productions nasales, elles se trouvent le plus souvent en position finale. D'après les auteurs, le passage du mode oral au mode nasal est dû à une diminution de l'énergie en fin de syllabe, et la tendance générale oral à nasal refléterait l'effet de la dominance des cadres. En d'autres termes, le résultat d'une forte énergie globale lors de la production de la syllabe serait un conduit vocal oral, alors que la nasalité serait la position par défaut lors de la diminution de l'énergie. Cette position de repos n'impliquerait aucun contrôle volontaire. Cette observation a été menée en comparaison avec une étude précédente de Coberly [3], qui donnait les résultats opposés. Il avait

observé plus de changements de modes dans le sens nasal à oral à l'intérieur d'une syllabe. Nous avons donc étudié ce point dans notre corpus.

Résultats

Le premier résultat que nous avons obtenu dans l'étude des nasales [11] est la proportion des consonnes nasales par rapport aux orales. Nous avons répertorié les consonnes orales et nasales dans la transcription complète du corpus. Nous n'avons pas comparé les nasales avec d'autres consonnes que les occlusives car la nasalité ne peut être produite dans une consonne non-occlusive. La première observation que nous pouvons déduire de ces résultats est que la proportion des nasales chez nos sujets diminue à partir de sept mois pour atteindre la proportion des nasales du Français (environ 10% des consonnes [23]), la langue maternelle à 11 mois. Ce fait pourrait nous permettre de considérer que le velum est pleinement contrôlé à ce stade. Ensuite, nous avons pu observer qu'à l'âge de 12 mois, la proportion des nasales a soudainement augmenté dans les productions de l'enfant mais ce phénomène peut être expliqué par un effet lexical puisqu'il coïncide avec l'apparition des premiers mots tels que "non" et "maman" dans le langage de l'enfant, bien identifiables sur la vidéo d'après le contexte.

Ensuite, nous avons étudié le contexte consonantique dans lequel apparaissent les consonnes nasales dans les séquences CVC. Nous avons observé un changement de mode oral vers nasal dès 6 mois à l'intérieur d'un cycle mandibulaire proto-syllabique, un indice évident pour un contrôle actif du velum. Nous avons observé un nombre important de production de séquences CVC à sept mois dans lesquelles nous avons repéré plus de changements du mode nasal au mode oral que l'inverse, contrairement aux résultats obtenus par Redford et al. [19]. Nous avons observé également les productions en contexte CCV afin de savoir si le même phénomène se reproduisait lorsqu'il n'y avait pas de voyelles entre les consonnes. En accord avec nos attentes, nous avons observé le même type de changement.

Ceci tend à montrer qu'il y aurait donc un mouvement actif du velum pendant l'oscillation mandibulaire, ce qui signifierait que le mouvement du velum ne serait pas impliqué dans la dominance des cadres. Il serait plutôt à considérer comme un paramètre indépendant dont le développement aboutirait progressivement à un meilleur contrôle des cadres proto-syllabiques.

2.3. Le contrôle de la coordination oro-laryngée

La coordination oro-laryngée, c'est-à-dire la coordination des gestes glottiques et supra-glottiques, est un principe de base de la communication humaine, peu étudié dans le babillage alors que la distinction voisé / non-voisé existe dans la plupart des langues du monde. Les études [6] menées sur cette coordination se sont principalement intéressées à la perception du VOT (Voice Onset Time) ou délai d'établissement du voisement qui est défini comme l'intervalle de temps entre le début des pulsations périodiques régulières glottiques et la détente de l'occlusion supra-glottique [13]. Il en ressort que les enfants perçoivent, dès 1

mois, une différence de VOT de 20 ms, à la frontière catégorielle adulte, mais nous ne savons pas à quel âge ils sont capables de reproduire cette différence qu'ils perçoivent. Nous avons donc entrepris une étude concernant la production du VOT et son évolution en effectuant des mesures du délai de voisement à partir des productions de deux enfants de 6 à 12 mois. Nous avons choisi des occlusives bilabiales, coronales et vélares bien saillantes [27] pour pouvoir effectuer nos mesures et nous avons appliqué la méthode de mesure du VOT de Lisker & Abramson [13].

Résultats

A partir des mesures de VOT effectuées dans cette étude [12] sur 400 plosives labiales, coronales et vélares en contexte vocalique ou pré-vocalique, nous avons obtenu deux catégories de VOT avec, dès 6 mois chez nos 2 sujets et pour tous les lieux d'articulation, une frontière similaire à celles des adultes francophones, soit des valeurs négatives de -95 à -10 ms et des valeurs positives de 5 ms à 65 ms, avec pour les deux sujets, une concentration des productions entre -60 et -20 ms pour les VOT négatifs et entre 15 et 50 ms pour les VOT positifs. Nous avons observé une nette séparation entre les VOT négatifs et positifs avec une frontière autour de 0 ms, tous lieux confondus, ce qui correspond aux productions adultes [13]. Mais les valeurs négatives sont, par contre, bien plus courtes chez l'enfant que chez l'adulte. Les enfants produisent des valeurs autour de -30 ms pour les occlusives voisées alors que les adultes francophones ont une moyenne à -130ms [26]. Ceci nous permet de supposer que les enfants auraient acquis la faculté de déclencher la vibration des cordes vocales après le relâchement de l'occlusion supra-glottique, mais qu'ils ne sont pas encore aptes à maîtriser le voisement dans l'occlusion à l'âge de 12 mois. Si les bébés présentent dès 6 mois et même avant des valeurs de VOT conformes à celles de leurs parents, il s'agit alors d'un contrôle très précoce de coordination glottique supra-glottique spécifique à la langue.

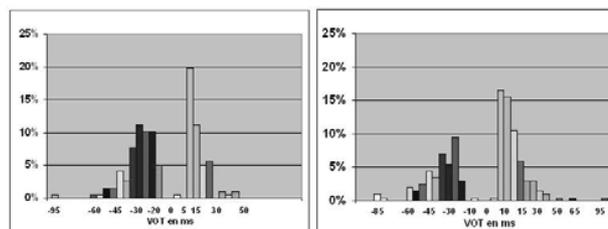


Figure 1 : valeurs du VOT lieux et âges confondus pour les sujets CB à gauche et NF à droite.

3. CONCLUSION

Les données articulatoires et acoustiques indiquent que les processus de nasalisation d'une voyelle ou d'une consonne impliquent des habiletés motrices très différentes. La position du velum pour une consonne nasale semble correspondre à une position par défaut, une position passive de l'articulateur. Chez l'adulte, le velum est la plupart du temps en position activement relevée, pour produire les sons oraux. Au contraire, les voyelles nasales requièrent un effort encore plus important, donc une plus grande habileté motrice pour obtenir une large ouverture du passage vélo-pharyngé.

Ceci expliquerait le fait que les consonnes nasales sont plus faciles à produire et donc plus fréquemment trouvées dans les premières vocalisations de l'enfant et c'est aussi pour cette raison que l'on ne retrouve pas de voyelles nasales au moins jusqu'à 12 mois.

De plus, nos observations nous laissent penser que le velum est déjà contrôlé à 6 mois car l'enfant est capable d'effectuer un mouvement du velum au cours de la production d'oscillations mandibulaires, mais il n'en est pas pour autant complètement contrôlé. Le premier stade de la maîtrise du velum pourrait se dater à 11 mois, au moment où les proportions nasales de l'enfant atteignent celles de l'adulte. La dernière étape du développement de ce contrôle serait atteinte lorsque le bébé est capable de produire des voyelles nasales et il semblerait que cela apparaisse plus tardivement dans le développement.

D'autre part, les mesures acoustiques du VOT révèlent qu'il existe, dès 6 mois, deux catégories distinctes voisée / non voisée pour les occlusives. Mais nous avons pu observer que les valeurs des enfants étaient, pour les VOT négatifs, bien inférieures à celles des adultes. Ceci révèle donc que les enfants n'arrivent pas encore à 12 mois à maintenir le voisement dans l'occlusion. On peut donc conclure que la première étape du contrôle est acquise mais qu'elle est peut-être due à un effet de contrainte aéro-dynamique. La seconde étape apparaîtra ultérieurement. En tout état de cause, ces contrôles représentent donc des éléments centraux dans la maîtrise progressive d'un contrôle global des cadres proto-syllabiques, chronologiquement situés entre l'émergence des cyclicités mandibulaires et le contrôle des segments du corps de la parole, à travers le contrôle des articulateurs portés que sont la langue et les lèvres. La maîtrise de ces habiletés motrices va permettre, à partir des cadres proto-syllabiques, le développement de séquences closant-vocant bien contrastées acoustiquement, c'est-à-dire, à proprement parler, de syllabes.

REMERCIEMENTS

Nous tenons à remercier Stefanie Brosda pour son travail d'enregistrement ainsi que pour avoir mis le corpus à notre disposition. Nous sommes aussi profondément reconnaissants envers Solange Rossato, Christian Abry, ainsi qu'envers les jeunes sujets et leurs familles.

BIBLIOGRAPHIE

- [1] S., Brosda, *Du babillage canonique à la naissance du contrôle des degrés de liberté des articulateurs*, Mémoire de Maîtrise, ICP, Université Stendhal, Grenoble III, 1998.
- [2] S., Brosda, *De la variation dans le babillage canonique : l'apprentissage sensori-moteur*, Mémoire de DEA Sciences Cognitives, ICP, Institut National Polytechnique de Grenoble, 1999.
- [3] M., Coberly, *Initial and final consonant preferences across languages, in babbling and in children's early speech*, PhD thesis, Linguistics, University of Colorado, 1985.
- [4] B.L. Davis, & P.F., MacNeilage, The articulatory basis of babbling, *Journal of Speech and Hearing Research*, vol.38, p. 1199-1211, 1995.
- [5] B.L. Davis & P.F., MacNeilage, An embodiment perspective on the acquisition of speech perception, *Phonetica*, n° 57, p. 229-241, 2000.
- [6] P., Eimas, E., Siqueland, P., Jusczyk, & K., Vigorito, Speech perception in infants, *Science*, n°171, p. 303-306, 1971.
- [7] A. L., Kinney, & P.F., MacNeilage, *Nasalization in vocalizations of pre-babbling infants : investigating an acoustic correlate*, *Speech Motor Control in Normal and Disordered Speech*, 4th International Speech Motor Conference, June 13-16, Nijmegen, the Netherlands, chap. 28, p.113-116, 2001.
- [8] A. L., Kinney, P.F., MacNeilage, B.L., Davis, "The role of nasalization in early vocal output", submitted to *Phonetica*.
- [9] F. J., Koopmans Van Beinum, "AMSTIVOC : testing and elaborating the Amsterdam system for transcription of infant vocalizations", *IFA Proceedings*, n° 23, p. 91-102, 1999.
- [11] C., Lalevée, A., Vilain, "Development of speech frame control : a longitudinal study of the oral/nasal control.", *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2003.
- [12] C., Lalevée, *Développement du contrôle de la production de parole, du cadre proto-syllabique vers la syllabe : suivi audio-visuel de 2 enfants de 6 à 12 mois.*, Mémoire de DEA Sciences du Langage, ICP, Université Stendhal Grenoble 3, 2003.
- [13] L., Lisker, & A.S., Abramson, "A cross-language study of voicing in initial stops, Acoustical measurements", *Word*, n° 20, p. 384-422, 1964.
- [14] J., Locke, The beginnings of phonology in the child, in Locke, J. (Ed.), *Phonological acquisition and change*, New York, Academic Press, chap.1, 1983.
- [15] P.F., MacNeilage, "The frame/content theory of evolution of speech production", *Behavioral and Brain Sciences*, n° 21, p. 499-546, 1998.
- [16] B., Macwhinney, *The CHILDES projet : tools for analysing talk, Third edition. Volume I : Transcription format and programs*. Mahwah, NJ : Lawrence Erlbaum Associates. ISBN 0-8058-2995-4, 2000a.
- [17] B., Macwhinney, *The CHILDES projet : tools for analysing talk, Third edition. Volume II : The database*. Mahwah, NJ : Lawrence Erlbaum Associates. ISBN 0-8058-3572-5, 2000b.
- [18] C.L., Matyear, P.F., Macneilage, & B.L., Davis, "Nasalization of vowels in nasal environments in babbling : evidence for frame dominance", *Phonetica*, n° 55, p. 1-17, 1998.
- [19] M.A., Redford, P.F., MacNeilage, & B.L., Davis, "Production constraints on utterance-final consonant characteristics in babbling", *Phonetica*, n°54, p.172-186, 1997.
- [20] S., Rossato, P., Badin, F., Bouaouini, "Velar movements in French : an articulatory and acoustical analysis of coarticulation", *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2003.
- [21] B. L., Smith, S., Brown-Sweeney, & C., Stoel-Gammon, "A quantitative analysis of reduplicated and variegated babbling", *First Language*, n° 9, p. 175-190, 1989.
- [22] M., Stefanuto, N., Vallée, Consonant systems : From universal trends to ontogenesis, *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1999.
- [23] J.P., Tubach, & L.J., Boë, *Un corpus de transcription phonétique (300.000 phones) : constitution et exploitation statistique*, Ecole Nationale Supérieure des Télécommunications, Paris, 1990.
- [24] F., Verin, *Etudes statistiques de données longitudinales du babillage : cyclicités, co-occurrences, closions*, Mémoire de Maîtrise, ICP, Université Stendhal, Grenoble III, 2001.
- [25] A., Vilain, C., Abry, P., Badin, S., Brosda, From idiosyncratic pure frames to variegated babbling : Evidence from articulatory modelling, *Proceedings of the 16th International Congress of Phonetic Sciences*, San Francisco, USA, August 1999.
- [26] W., Serniclaes, *Etude expérimentale de la perception du trait de voisement des occlusives du Français*, Thèse – spécialité Sciences Psychologiques, Institut de phonétique, Université Libre de Bruxelles, 1987.
- [27] U., Sundberg & F., Lacerda, Voice onset time in speech to infants and adults, *Phonetica*, n° 56, p. 186-199, 1999.