

Le cadre de la parole et le cadre du signe : un rendez-vous développemental

Virginie Ducey & Christian Abry

Institut de la Communication Parlée UMR CNRS 5009
INPG/Université Stendhal, DU, 1180 avenue Centrale BP25 – 38040 Grenoble cedex 9 France
Tél : +33(0)4 76 82 41 28 – Fax : +33(0)4 76 82 43 35
Mèl : ducey@icp.inpg.fr / abry@icp.inpg.fr

ABSTRACT

We raise the question of the primitives of human communication, especially with regard to their development. This question seems crucial if one wants to understand and study the devices that are required in the emergence of language in the infant, through his/her behaviour in speech and gesture.

We start from a non a priori predictable *rendez-vous* discovered quite recently : the co-emergence of a basic speech control, namely *coarticulation*, with the first word(s) just after one year. Consequently we ask three empirical questions : (i) What is needed in speech control to get coarticulation ? (ii) What is needed in sign control to get a word ? (iii) What is the metric which can allow cross-talk and integration of sign and sound ?

The natural time-window for our investigation will naturally be set about the second half of the first year.

1. INTRODUCTION

Nous posons ici la question des primitives de la communication humaine : comment se développent-elles chez les humains, voire chez les primates non-humains ? Ces primitives nous semblent particulièrement importantes si l'on veut essayer de comprendre et étudier les mécanismes qui entrent en jeu dans l'émergence du langage chez l'enfant, à travers ses comportements en paroles et gestes.

Pour cela, nous partons de l'hypothèse d'un « rendez-vous développemental » entre ce que nous nommons le *cadre du signe*, et le *cadre de la parole*. Cette hypothèse repose sur une découverte récente [5], à savoir la co-émergence du phénomène de coarticulation — qui constitue le contrôle de base de la parole — avec les premiers mots chez l'enfant, juste à la fin de la première année.

Ceci nous amène à aborder plusieurs questions empiriques : (i) de quels mécanismes l'enfant a-t-il besoin pour arriver à l'étape de la coarticulation pour le contrôle de la parole; (ii) de quels mécanismes a-t-il besoin pour arriver au mot, dans le contrôle du signe; et enfin (iii) quelle est la métrique qui peut permettre le dialogue et l'intégration entre le signe et la parole ? Nous allons donc

essayer dans cet article de présenter ces différents mécanismes fondamentaux et leur intégration.

2. LE CADRE DE LA PAROLE (LA MANDIBULE SYLLABIQUE)

Nous allons brièvement résumer les différentes étapes requises pour le développement du cadre de la parole (d'après [1]).

Nous partons de la théorie Cadre/Contenu (Frame/Content) de MacNeilage, qui est une théorie du contrôle de la parole à la fois dans son développement et dans son évolution [2]. MacNeilage argumente que la parole humaine diffère de la communication vocale de tout autre mammifère. Similairement à d'autres espèces animales, l'être humain utilise des séquences d'ouverture et fermeture du conduit vocal dans le but de communiquer. C'est ce que MacNeilage nomme le *cadre*. Néanmoins, nous nous distinguons des autres espèces par le fait que nous ajoutons à ce mouvement d'alternance une modulation articulatoire, ce qui correspond selon l'auteur au *contenu*. Afin d'éclairer le chemin évolutif de la parole humaine, MacNeilage propose de s'interroger sur l'ontogenèse de la parole, si l'on admet qu'elle peut récapituler, au moins en partie, la phylogénèse.

MacNeilage note qu'à l'âge de 5 mois, les bébés sont déjà capables de jouer avec leur mâchoire (*jaw wags*). Mais, vers l'âge de sept mois, émerge chez l'enfant ce que l'on appelle le *babillage canonique*, qui correspond typiquement à des productions du type [bababa...] ou [dadada...] [2]. [1] ont pu démontrer sur différents modèles articulatoires, respectant les différences individuelles, que la nature du contact — labial [b] et/ou coronal [d] — dépendait uniquement de la configuration du conduit vocal de l'individu. En d'autres termes, le contact n'est pas du tout programmé : il dépend uniquement du mouvement mandibulaire. Notons que ce babillage apparaît aussi chez l'enfant sourd, mais avec un certain retard. Du point de vue du contrôle moteur, MacNeilage propose que ce babillage canonique correspondant à l'apparition du *cadre* de la parole, soit un contrôle des oscillations mandibulaires qui dépend d'un système d'initiation motrice, et MacNeilage a proposé l'aire motrice supplémentaire (SMA). C'est une candidature neurale que [3] ont pu soutenir en proposant que les patients comme « tan, tan » de Broca n'aient plus

que ce contrôle de la parole. Selon [1], le contrôle du *cadre* est bien celui de la *porteuse* de la parole, c'est-à-dire le contrôle de cet *effecteur proximal* qu'est la mandibule. Ce cadre ne sera rempli que plus tard par un *contenu segmental*. Toujours selon [1], il s'agira alors du contrôle des articulateurs *portés*, ou *effecteurs distaux* (langue, lèvres...).

Cette différence entre contrôle proximal et contrôle distal est bien mise en évidence dans l'enregistrement ci-contre d'une fillette de 8 mois par [4] (Figure 1). Sur le tracé du haut, les mouvements enregistrés pour la lèvre supérieure sont passifs, uniquement dus à la poussée de la lèvre inférieure, qui elle-même est portée par la mandibule. Par comparaison, chez l'adulte (tracé du bas), on voit : (i) que la lèvre supérieure résiste activement à la poussée de la lèvre inférieure ; (ii) et que cette même lèvre supérieure se relève activement en coordination avec l'abaissement de la lèvre inférieure pour produire l'ouverture de la voyelle.

On peut observer typiquement, dans le cours du développement, pour les mouvements de type [bababa], que la lèvre supérieure devient de moins en moins compliant à la poussée de la lèvre inférieure portée par la mandibule. Ce qui aboutira finalement à un contrôle des contacts autonome.

À l'âge de 13 mois naît la coarticulation dans le babillage canonique, en même temps que naissent les premiers mots eux aussi coarticulés (d'après notre lecture des données de [5]). Dans une séquence [aba], la langue en montant de [a] vers [b], ne monte pas autant qu'elle ne monterait si elle était passivement portée par la mandibule, laquelle porte la lèvre inférieure à la rencontre de la lèvre supérieure pour effectuer la closure de [b] : autrement dit, la langue résiste par un contrôle musculaire (hyoglosse) à la montée mandibulaire. C'est en fait la définition même du contrôle de la coarticulation, ou comment produire un [a] pendant que la bouche est fermée par la production d'un [b]. La possibilité de préparer la position du [a] pendant la fermeture permettra d'avoir des formants caractéristiques de la voyelle très vite atteints dès l'ouverture du conduit vocal.

Tirant profit de cette coïncidence surprenante — notre « rendez-vous développemental » — entre la coarticulation et les premiers mots, et sans préjuger d'une explication possible, nous pouvons maintenant nous interroger sur le phénomène qui lie l'émergence du contrôle de la parole au contrôle du signe.

3. LE CADRE DU SIGNE (DE L'ŒIL À L'INDEX)

S'agissant du contrôle de la production des premiers mots, nous aurons avantage à ne pas le distinguer dans un premier temps du contrôle d'un autre signe, le signe de la langue des signes. Dans la recherche de ce *cadre du signe*, il nous semble nécessaire d'intégrer les principaux mécanismes utiles au contrôle de la sémantique des premiers mots comme des premiers signes.

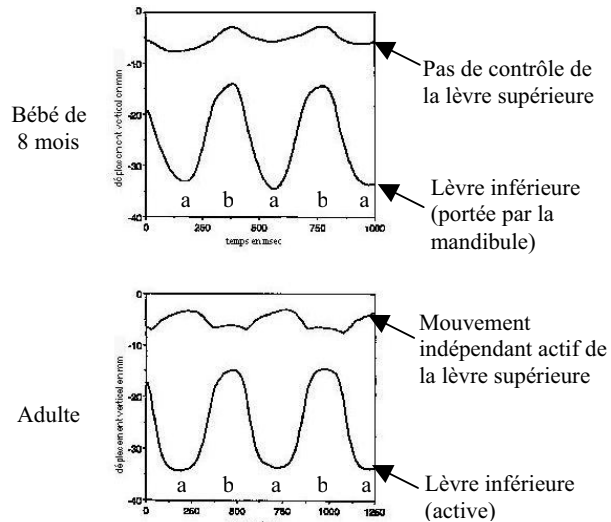


Figure 1 : Munhall et Jones (1998, p.525)

Nous nous inspirons de la notion d'« objectitude » (*objecthood*), qui fournit une base pour l'apprentissage du langage. L'*indexation* des *agents* (qui semblent être les mieux reconnus chez l'enfant) des *objets* et des *événements*, constitue une théorie permettant d'ancrer la capacité de suivi de une à trois cibles dans l'*analyse de scène*, comme le proposent, entre autres [6]. Cette *indexation* se fait en premier lieu par l'œil, bien avant que le pointage de l'index ne se mette en place chez l'enfant, vers 9-12 mois.

Elle se réalise en parallèle avec l'émergence de différents modules d'attention de [7]. Ces modules d'attention sont nécessaires pour le développement des trois grands systèmes sémantiques de suivi d'agent/objet/événement également fondamentaux pour le cadre du signe. Ces trois systèmes peuvent être considérés comme de véritables racines pour l'acquisition lexico-grammaticale du langage.

Il est reconnu que les enfants individualisent des *événements*, *objets*, *agents* en mémorisant non pas leurs propriétés, mais d'abord leurs localisations simultanées et leurs permanences, y compris lorsque ces objets sont cachés derrière un écran. Nous avons repris (d'après [6]) pour cette fonction de localisation sémantique les acquis sur l'organisation des deux voies du cerveau, « sémantique » et « pragmatique », selon les qualifications de Marc Jeannerod, qui sont respectivement la voie ventrale « What », et la voie dorsale « Where ».

Par le système pragmatique, la voie que nous appelons « Where/There » (toute interrogation mérite réponse), l'enfant établit une assignation de l'objet à son cadre spatio-temporel. Si l'on montre à un enfant de 10 mois un jouet canard et un jouet voiture qui apparaissent et disparaissent, un jouet à la fois, derrière un écran, ces enfants vont s'attendre à un seul objet lorsque l'écran est relevé. Par contre, quand on leur montre le canard et la voiture apparaissant et disparaissant simultanément

derrière cet écran, les enfants de 10 mois s'attendent à deux objets derrière l'écran. Ainsi les enfants paraissent bien individuer les objets par l'information spatio-temporelle.

Plus tard, après 10 mois, l'enfant peut se passer de l'information spatio-temporelle. En effet à la place d'individuer un objet par cette information, l'enfant va individuer un objet et l'identifier par l'information de traits, par exemple par sa forme. Ainsi, dans l'expérience précédente, on montre que, même si le canard et la voiture ne sont jamais vus simultanément, l'enfant s'attendra, avec des formes différentes, à deux objets derrière l'écran. La forme est l'une des premières propriétés à entrer dans ce système sémantique cérébral ventral « What/That ». Ce n'est que vers 12 mois, que les enfants montreront la capacité d'individuer les objets par leur propriété de couleur. Avant un an, si deux canards de différentes couleurs ne sont jamais vus simultanément, l'enfant ne s'attendra pas à deux canards derrière l'écran.

On peut montrer que le premier système « Where/There » n'est pas absolument prioritaire. De fait, il est préempté par un autre système pragmatico-sémantique, le système « When/Then », qui se focalise sur l'interrogation et la réponse temporelle. [8] ont ainsi pu mettre en évidence cette influence du système « When » sur le système « Where » avec une expérience, dans laquelle deux cibles visuelles (deux boules), partant du haut de l'écran, se croisent en X. Elles peuvent être perçues effectivement comme se croisant ou bien comme se heurtant puis se séparant, ceci à condition qu'un événement se produise autour de leur point de rencontre. Ce peut être un flash, un bip ou un retour d'effort tactile : peu importe la modalité. Les sujets suivent le mouvement grâce au système « Where », mais la présence du système « When » peut changer la trajectoire du suivi.

Par l'œil, et puis par l'index, on assiste ainsi à la mise en place des systèmes « When, Where, What ».

4. QUAND LE CADRE DE LA PAROLE S'INTÈGRE DANS LE CADRE DU SIGNE (LE BRAS DEICTIQUE)

Quelle évidence avons-nous que le cadre du signe intègre le cadre de la parole ? Nous nous appuyerons sur des données récentes, celles de [9], enregistrant des mouvements de la main chez des bébés entendant nés de parents sourds signant, quand ils étaient âgés de 6, 10 et 12 mois. Les analyses des signaux de mouvements donnés par les capteurs posés sur les membres supérieurs des bébés, révèlent que ceux qui sont exposés à la langue des signes, montrent pour les mouvements rythmiques, une manière significativement différente d'utiliser l'activité de la main par rapport à ceux exposés à l'oral.

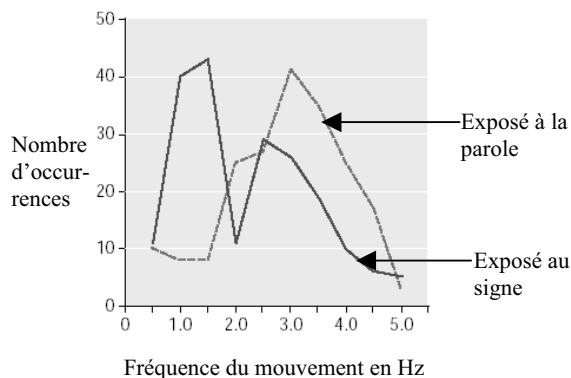


Figure 2 : Petitto et al. (2001, p.35)

La figure 2, tirée de [9] affiche pour nous un résultat majeur. Elle donne la distribution des fréquences des mouvements rythmiques de la main mesurées pour l'activité des bébés exposés aux signes (ligne pleine), comparée à la distribution pour les bébés exposés seulement à la parole (ligne tiretée). Petitto et al. remarquent que seuls les bébés exposés aux signes ont une distribution bi-modale des fréquences de mouvement : le premier mode (pic de gauche sur la figure 2) est à 1.5 Hz, et le second mode (pic de droite sur la figure 2) à 2.5 Hz. Ils comparent alors ces résultats avec les fréquences de mouvement des mains des bébés exposés à la parole, qui elles sont unimodales, à 3 Hz. Au regard de cette comparaison, le pic le plus élevé (2.5 Hz) produit par les bébés exposés aux signes n'est pas significativement différent du pic à 3 Hz des bébés exposés à la parole. Par contre, évidemment, le mode à 1.5 Hz est propre aux enfants exposés aux signes.

Ce qui semble en outre spécifique chez ces bébés exposés à la langue des signes, c'est que l'activité de la main s'effectue dans un espace bien restreint, correspondant à l'espace utilisé par le signe des adultes signeurs. A l'inverse, les bébés exposés à la parole, voient la majorité de leur activité des mains se situer hors de l'espace de la langue des signes.

Cette étude nous semble particulièrement pertinente pour notre questionnement, car elle nous amène — de la même manière que nous avons cherché le « pas de la parole » dans le rythme syllabique — à rechercher le « pas du signe ». L'unité métrique connue des linguistes qu'est le pied nous fournit un gabarit commun pour intégrer la syllabe dans ce pas du signe. Notons que ce pied (*foot*) est une unité métrique accentuelle (un seul accent par pied) qui, du point de vue du contrôle moteur, rend compte de la tendance à l'isochronie (comme en anglais dans les mots *ease*, *easy*, *easily*). Notre hypothèse repose donc sur le fait que chez les enfants exposés aux parents sourds signant, il existe un mouvement plus lent des mains, un mode de 1.5 Hz, sachant qu'à ce mode, il s'agit de proto-signes. A ce stade, on peut remarquer que lorsque les bébés produisent leurs premiers mots, ces « mots prosodiques » sont de manière variable, d'une ou de deux syllabes. En fait, notre idée est que pour avoir un signe-

mot en parole, il faut arriver à intégrer deux cycles de la mandibule (le babillage canonique peut se faire à environ 3 Hz, soit 300-350 ms) dans un pied (1.5 Hz, soit 600-700 ms). Ce que l'enfant entendant ou sourd finit par faire avec la mandibule (cadre de la parole), il l'intègre dans le bras qui pointe, l'index de la main tendu, soit dans le *cadre* du signe, en coordination avec la prosodie vocale : autrement dit le mot.

5. CONCLUSION

Nous avons présenté ici notre dispositif général de recherche, en reprenant les mécanismes fondamentaux dans la mise en place de ces deux flux, le *cadre du signe* et le *cadre de la parole*, qui aboutissent en conjonction aux premiers mots chez l'enfant. Il nous a semblé en effet nécessaire de nous attarder, dans un premier temps, sur l'aspect développemental de ces cadres, que cela soit dans l'exploration du contrôle du conduit vocal pour le cadre de la parole ou l'exploration des mécanismes d'attention et des systèmes sémantiques pour le cadre du signe. Nous avons besoin de ces ancrages de contrôle des segments corporels pour comprendre comment s'établit ensuite le dialogue entre ces deux flux, et comment le cadre du signe arrive à intégrer le cadre de la parole.

Des études empiriques nous permettront de rendre compte de cette intégration, par l'intermédiaire d'une observation que nous menons sur un corpus audio-visuel réalisé longitudinalement sur 6 enfants de 7 à 18 mois. Ce corpus nous permettra de relever les événements de pointage de l'index, tout particulièrement la phase de détente du bras ou « stroke », qui se définit depuis l'initiation du mouvement du bras jusqu'à la position du bras complètement tendu. C'est le cadre qui sert de pas au contrôle du signe, en coordination avec les unités de la parole, les segments contenus dans les pas syllabiques. Nous pourrions ainsi réaliser des mesures et une analyse quantitative pour soutenir l'hypothèse que l'on peut intégrer une à deux syllabes dans un « stroke » de quelques 600 ms, confirmant ainsi l'intégration du cadre de la parole dans le cadre du signe.

Remerciements :

Stefanie Brosda pour le corpus longitudinal de 7 à 18 mois ; Anne Vilain pour la partie *cadre de la parole* ; Alain Arnal pour le montage de la démonstration du test de Shimojo et al. que nous avons baptisé le test « When ? Where ? What ? ».

6. BIBLIOGRAPHIE

- [1] A. Vilain, C. Abry, P. Badin. Coproduction strategies in French VCVs : Confronting Öhman's model with adult and developmental articulatory data. *5th Seminar on Speech Production: Models and data, Munich, Kloster Seon, Bavaria*, 81-84, 2000.
- [2] P. MacNeilage. The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21/4, 499-511, 1998.
- [3] C. Abry, M. Stefanuto, A. Vilain, R. Laboissière. What can the utterance "tan, tan" of Broca's patient Leborgne tell us about the hypothesis of an emergent "babble-syllable" downloaded by SMA ? In *Phonetics, Phonology and Cognition*, J. Durand & B. Laks (Eds.), Oxford University, 226-243, 2002.
- [4] K.G. Munhall, J.A. Jones. Articulatory evidence for syllabic structure. *Behavioral and Brain Sciences*, 21/4, 524-525, 1998.
- [5] H. Sussman, C. Duder, E. Dalston, A. Cacciatore, Acoustic analysis of the development of CV coarticulation: a case study. *Journal of Speech, Language, and Hearing Research*, 42, 1080-1096, 1999.
- [6] A.M. Leslie, F. Xu, P. Tremoulet, B. Scholl. Indexing and the object concept: Developing "What" and "Where" systems. *Trends in Cognitive Sciences*, vol. 2, n°1, 10-18, 1998.
- [7] S. Baron-Cohen. *Mindblindness*. MIT Press, 1995.
- [8] S. Shimojo, C. Scheier, R. Nijhawan, L. Shams, Y. Kamitani, K. Watanabe. Beyond perceptual modality: Auditory effects on visual perception. *Acoustic Science & Technics*, 61-67, 2001.
- [9] L.A. Petitto, S. Holowka, L.E Sergio, D. Ostry. Language rhythms in baby hand movements. *Nature*, vol. 413, 35-36, 2001.