Development of speech frame control: a longitudinal study of the oral/nasal control.

Lalevée, C. & **Vilain, A.**
Institut de la Communication Parlée(ICP) - CNRS UMR 5009
Université Stendhal, BP 25, Grenoble cedex 9, France
e-mail: lalevee@icp.inpg.fr, *avilain@icp.inpg.fr*

Abstract

The unfolding of events in speech ontogeny can be described as the progressive mastering of the degrees of freedom of speech articulators, a process that is to be related to training, imitation, and neural maturation in the child's brain. In this framework, we have shown in previous studies, that canonical babbling, a major step in language development, can indeed be related to the sudden emergence of control over the carrier articulator, i.e. the mandible, in the absence of any other voluntary control over the other speech articulators such as the velum, the lips and the tongue. However if the mastering of mandibular oscillations can be described as the basic carrier structure in speech, this protosyllabic frame will not be adapted to the needs of human adults 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. One element in this global control is that of the velum, which enables to have a fully oral vocal tract, therefore to produce salient consonant-vowel sequences. We have studied the evolution of this control in speech development from an audio-visual digital video corpus of 6 French children. We have been able to show that there is an active control of the velum as soon as 6 months, evidenced by lowering and raising movements during an utterance, and that the proportion of nasal sounds reaches adult norms about 11 months, indicating that a first stage in the mastering of the velum is by that time reached, the next one being the control of nasal vowels.

1. Introduction

1.1. Development of frame control

Babbling, as the first appearance on the stage of speech development of series of open-close alternations of the vocal tract, has been the focus of a number of studies for many years. It has been considered as the emergence of proto-syllables: the first time when a baby is able to coordinate glottal gestures with supra-glottic rhythmic movements, i.e. the basic structure of human communication. In his "Frames, then Content" theory of speech communication, MacNeilage [11] has put forward the idea that these proto-syllabic frames are due to the unique oscillation of of the mandible in the absence of any other control over carried articulators. Many studies after this founding paper have endeavoured to test the weight of mandibular oscillations, and the extent to which they alone can explain the vocal products of babbling. The correlation between the places of articulation of consonants and vowels within one babbling "syllable" has been investigated and taken as an argument for the non-independance of these two parts of a protosyllabic sequence, therefore an evidence that these two parts are to be attributed to the same movement of one single articulator, the mandibular carrier (Davis et al. [5], Matyear et al. [14], Vérin [21]). Another study [Vilain et al., 22] has tested the possibility that mandibular oscillations could actually be the source of babbling products, by reproducing them with articulatory models on which all degrees of freedom except for the mandible had been frozen. The results supported the theory: their pool of models, in spite of being too small to constitute a representative sample, was sufficient to yield the two main consonantal place types observed in babbling databases.

However if the controlling of mandibular oscillations can indeed be described as the basic carrier structure in speech, this protosyllabic frame will not be adapted to the needs of human adults 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. One element in this global control is that of the velum, which enables to have a fully oral vocal tract, therefore to produce salient consonant-vowel sequences. We have run a first study on the development of the control of an oral vocal tract, with the mastering of the velum, evidenced by the evolution of nasal consonants in pre-speech babbling.

1.2. The oral/nasal control: a major step in speech development

Nasal consonants represent 14.6% of the consonantal sounds of the world's languages (cf the UPSID 451 database, Stefanuto [18]), nasal vowels are 22.4% of the vowels (Vallée, [20]). The explanation we can put forward for such a small proportion is the relatively bad auditory

ISBN 1-876346-48-5 © 2003 UAB

salience of these sounds in a syllable. A sequence consisting of a nasal consonant followed by a vowel has little auditory contrast, since the nasal consonant, as well as the vowel, shows a large amplitude and a formant structure. Such a sequence hardly shows the non-sonorant – sonorant pattern that characterizes a syllable in the majority of cases in the world's languages. Moreover nasals' places of articulation are poorly identifiable: languages do not exploit many different places of articulation.

These sounds are produced with the velum raised against the pharyngeal wall, that is an active movement of the velum. The question is then: is a closed velopharyngeal port the default position of the velum from birth ? A number of studies [9, 17] have shown that baby vocalizations contain a large proportion of nasal sounds. Then how and when is the control of this particular velum gesture acquired by the baby ? We have studied the evolution of the proportion of nasals in babbling, to determine when the baby reaches adult-like norms. The evolution of nasality was evaluated from the proportion of nasal consonants.

1.3. Different controls for nasal consonants and vowels ?

Opening the nasal tract by way of lowering the velum has very different consequences, whether the vocal tract is open or closed. If the vocal tract is closed as in occlusives, the nasal fossae will act as a resonator and induce the appearance of formants, therefore a dramatic change in the acoustic structure of the consonant. Thus the difference between the two positions of the velum in a consonant yields very obvious and unambiguous cues on the speech signal. Starting from an open vocal tract to lower the velum, i.e. changing from an oral to a nasal vowel does not modify radically the acoustic nature of the sound, but has rather slighter effects, such as widening the bandwidth of the first formant, lessening the amplitude of this first formant, lowering the frequency of the second formant...Actually, the detection of reliable acoustic correlates of nasality in vowels still represents a stumbling block for acoustic researches. Spectral shape for example was once supposed to be a good correlate of nasality, but it proved unreliable (Kinney, [7]). Detecting nasal consonants rather than nasal vowels is therefore more reliable, and all the more in baby vocalizations.

A second argument in the favour of studying nasality in consonants is to be found in Rossato et al. [16]. They have studied the movements of the velum in a French speaker through electro-magnetic articulometry measurements, and particularly the distribution of velum vertical positions for oral versus nasal vowels, and oral versus nasal consonants. Oral vowels are here to be understood as parts of the French phonological oral category, and acoustically analyzed as oral. What comes out of this study is that the difference from an oral to a nasal consonant can be realised by a mean 2 mm movement of the velum, whereas a nasal vowel needs a much more widely open nasal tract and so a larger downward movement of the velum (mean difference between the high and low positions of the velum: 7 mm). This implies that changing from the oral to the nasal mode in consonants represents a small effort, needing little energy, while the change in vowels is an effortful gesture. Moreover it appears that an oral vowel can even be produced with a small aperture of the nasal tract, especially for low vowels. In other words, a small opening of the velopharyngeal port will not yield radical acoustic changes in an oral vowel whereas it will change radically the acoustic structure of a consonant, and a [mama] sequence can be produced with a stable slightly lowered position of the velum.

In summary, if the slightly open configuration of nasal consonants may be considered as a default position of the velum in the early stages of development, the low position for nasal vowels may imply a complex control of the velum, that should develop much later in ontogeny. This sheds light on the studies that should be carried on about nasality in babbling vocalizations: no strong effect is to be expected in the acoustic structure of vowels until the time when the baby is able to control a large opening of the velopharyngeal port, while nasal consonants may show the evolution of a first stage in the development of this control.

1.4. Evidence for an active control

Redford et al. [15] have carried on a study about production constraints on utterance-final consonants in babbling, from the beginning of babbling to the age of 12 months. In particular, they focussed on changes in oral vs nasal manner within one utterance, and found more changes from oral to nasal, that is more nasals in final position. The interpretation they gave to this result is that the change from oral to nasal would be merely due to a decrease in the amount of energy at the end of an utterance. In other words, an oral vocal tract would result from strong global energy given to the production of an utterance, while nasality would be a default position when the energy decreases, a relaxation position, implying no voluntary control. This observation had been made as a comparison with a former study by Coberly [4], that gave the opposite results: more changes from nasal to oral manner within one utterance. We studied this point in our corpus.

1.5. A control developing on auditory stimulation

Many studies have shown that infants with hearing loss produce a larger proportion of nasal consonants than normal hearing infants. This suggests that the infant must be able to hear that the adult language is predominantly non nasal in order to motivate changing from the resting position of an open velopharyngeal port.

McCaffrey et al. [10] have studied the evolution of nasality in a 25-month profoundly hearing-impaired infant receiving a multichannel cochlear implant. They showed that pre-implant productions were around 80% nasal consonants and that this proportion reduced dramatically about seven months after the implant, to reach the same proportions as normally hearing infants. This suggests that the developing infant needs access to auditory information to stimulate the oral production mode. They conclude that if normal access to extrinsic and intrinsic information is available, an infant will be capable of achieving the low frequency of nasality in the ambient language within seven months.

2. Corpus and method

2.1 An audio-visual corpus

The corpus we have worked on was recorded by S. Brosda [3] and consists in audio-visual recordings (VHS + DAT, then DV) of 6 French children aged 6 to 15 months. The infants have been recorded in their family environment during 45 minutes every fortnight, from the date when the parents signalled the begin of the babbling until 9 months later.

The present study bears on the data of one child from the age of 6 to 12 months, from babbling onset to the first words. It consists in 10 sessions. The acquisition of the corpus was made by S. Brosda. The audio data were digitally acquired at a sampling rate of 22.050 Hz, and the child's productions were compiled in a separate file, each vocalization separated from the next by a 2 second silence. The vocalizations that were retained are productions that corresponded to the babbling criteria [8] (no burps, laughs, coughs, or raspberries) .

2.2 The transcription

The corpus has been transcribed through successive analysis by two phonetician experts, using CLAN (Computerized Language ANalysis), a program from the CHILDES Database (The CHIld Language Data Exchange System) (MacWhinney [12,13]). The transcription method is inspired by the AMSTIVOC system (Amsterdam System for Transcription of Infant VOCalizations, Koopmans van Beinum [8], Brosda [2]). It is based on a source-filter description of infant productions, and describes the phonation type and the articulatory gestures employed rather than attributing phonetic symbols to productions. In this method, the segmentation unit is one respiration cycle.

The overall structure of the productions is described with: number of rhythmic unities, number of visible jaw wags, number of silent jaw wags, phonation type throughout the utterance, serial position of the constrictions if any, and number of constrictions within the utterance.

Then each rhythmic unity is detailed with (i) the consonantal sound: place of articulation, the manner of constriction, voicing, type of change from one closant to the next, (ii) the vocalic sound: vocant quality (defined as an area of the F1/F2 vocalic space divided in twelve parts), type of change from one vocant to the next, extra vocalic qualities, and (iii) global place of articulation of the rhythmic unity. Finally the transcription includes the number of CV-like unities and the utterance type in term of reduplication and variegation.

The transcription of the audio files was complemented with the video recordings, and based on acoustic analysis with the PRAAT speech processing software. We used the VLAM model [1], a model of the growth of the vocal tract, to help transcribing the vowels, by figuring the mean vocalic space for a child of that age.

3. Results

The first outcome of the study is the number of nasal consonant-like sounds, as a proportion to oral occlusive consonants for the child CB (Figure 1). We are not comparing nasals with consonants other than occlusives, since nasality would not be detected in a non-occlusive consonant. The first observation we can draw is that the proportion of nasals keeps lessening from seven months on. The proportion is 52.94% at 7 months, and 16.86% at 11 months. At this age, the baby has reached the proportion of nasals in French, her native language (Tubach, [19]), a fact that we may consider as an evidence for a fully mastered velum. It is to be noted that at the age of 12 months, the proportions of nasal sounds suddenly increase greatly in the infant's productions. This phenomenon is to be explained by a lexical effect, with the appearance of words such as "non" (no) and "maman" (mummy) in the child's language, and lead us to remove this data from our study.

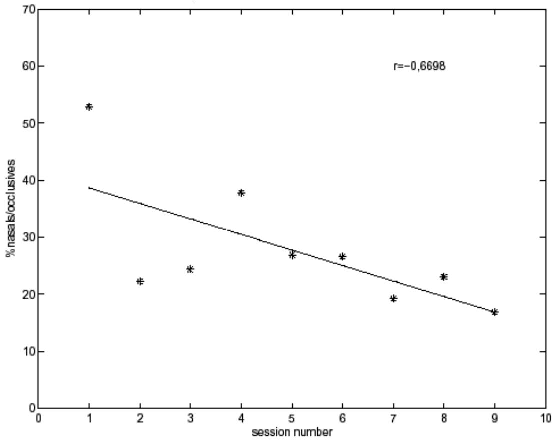

Figure 1: proportion of nasal consonants against oral occlusives, for child CB, from 6 to 11 months

Second, we studied the contexts in which nasal closants appear in CVC syllable-like sequences. We observed from 6 months the changes in manner from oral to nasal within a mandibular protosyllabic cycle, an evidence of an active control of the velum. We observed an important production of CVC sequences from seven months on and in these CVCs we observed more changes from nasals to orals than changes from orals to nasals, contrary to the Redford et al. results [14]. We observed the productions in context CCV to see if the phenomenon was the same when there were no vowel between the consonants. The same type of changes as previously was observed. This stands as an element to attest an active movement of the velum during mandible oscillations, which would mean that velar movements are not implicated in frame dominance, and should rather be considered as an independent parameter, whose development will eventually ensure better frame control.

ISBN 1-876346-48-5 © 2003 UAB

4 Conclusion

Articulatory and acoustic data indicate that nasalizing a vowel or a consonant implies very different motor skills. The velum position in a nasal consonant seems to correspond to a default, passive position of this articulator, which is actually very close to the position required for oral sounds, while nasal vowels require a greater effort, therefore better motor skills, to widely open the velopharyngeal port. This explains the fact that nasal consonants are easily and frequently found in early infant vocalizations, whereas no strong evidence is found for nasality in vowels, at least until 12 months of age.

Our observations lead us to think that the velum is already controlled at 6 months, because the child can move her velum during the production of sound and movement of the mandible, but is not yet completely mastered. The mastering of the velum should be dated for this child about 11 months, when she reaches adult-like proportions of nasals against oral consonants. The final stage in the development of this control will be reached when the baby is able to produce fully nasal vowels, and this is to be tracked later in the development.

Data from hearing-impaired children suddenly accessing auditory information through multichannel cochlear implantation, support the idea that this control develops on the basis of the perceptuo-motor circuit. These data indicate that the production of oral sounds is stimulated by auditory information, and it would be interesting to know how far the development of this control is guided by these auditory information, by comparing the developments of children from different native languages, particularly children whose mother tongue does not contain any nasal, for example.

The control of a vocal tract with raised velum stands as a significant stage in speech development chronologically situated between the emergence of the mandibular cyclicities and the development of the segmental content through accurate control of the carried articulators (= lips, tongue). This control is a major index of the progressive mastering of fully controlled proto-syllabic frames, also evidenced in the development of efficient oro-laryngeal coordination. These motor skills will allow the development of acoustically salient CV sequences, that is actual syllables, from protosyllabic frames.

Acknowledgements

We would like to thank Stefanie Brosda for the huge amount of data she has collected and put at our disposal. We are also deeply grateful to Solange Rossato, to Christian Abry, and to our young subjects and their families.

References

[1] Boë, L.-J, Maeda, S.. Modélisation de la croissance du conduit vocal. Journées d'Études Linguistiques "*La Voyelle dans tous ses états*". Nantes, 98-105, 1998.

[2] Brosda, S., *Du babillage canonique à la naissance du contrôle des degrés de liberté des articulateurs*, Mémoire de Maîtrise, ICP, Université Stendhal, Grenoble III, 1998.

[3] Brosda, S., *De la variation dans le babillage canonique: l'apprentissage sensori-moteur*, Mémoire de DEA Sciences Cognitives, ICP, Institut National Polytechnique de Grenoble, 1999.

[4] Coberly, M., *Initial and final consonant preferences across languages, in babbling and in children's early speech*, PhD thesis, Linguistics, University of Colorado, 1985.

[5] Davis, B.L. & MacNeilage, P.F., "An embodiment perspective on the acquisition of speech perception", *Phonetica*, 2000, n° 57, p. 229-241.

[6] Kinney, A. L., & MacNeilage, P.F., "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, 2001, Nijmegen, the Netherlands, chap. 28, p.113-116.

[7] Kinney, A. L. MacNeilage, P.F., Davis, B.L., "The role of nasalization in early vocal output", submitted to *Phonetica*.

[8] Koopmans Van Beinum, F. J., "AMSTIVOC: testing and elaborating the Amsterdam system for transcription of infant vocalizations", *IFA Proceedings*, n° 23, p. 91-102, 1999.

[9] Locke, J. The beginnings of phonology in the child, in Locke, J. (Ed.), *Phonological acquisition and change*, New York, Academic Press, chap.1, 1983.

[10] McCaffrey, H.A., Davis, B.L., MacNeilage, P.F. & Von Hapsburg, D., "Multichannel cochlear implantation and the organization of early speech", *The Volta Review*, vol. 101(1), p.5-29, 2000.

[11] MacNeilage, P.F., "The frame/content theory of evolution of speech production", *Behavioral and Brain Sciences*, n° 21, p. 499-546, 1998.

[12] Macwhinney, B., *The CHILDES project: tools for analysing talk, Third edition. Volume I: Transcription format and programs*. Mahwah, NJ: Lawrence Erlbaum Associates. ISBN 0-8058-2995-4, 2000.

[13] Macwhinney, B., *The CHILDES project: tools for analysing talk, Third edition.* Volume II: The database. Mahwah, NJ: Lawrence Erlbaum Associates. ISBN 0-8058-3572-5, 2000.

[14] Matyear, C.L., Macneilage, P.F. & Davis, B.L., "Nasalization of vowels in nasal environments in babbling: evidence for frame dominance", Phonetica, n° 55, p. 1-17, 1998.

[15] Redford, M.A., MacNeilage, P.F. & Davis, B.L., "Production constraints on utterance-final consonant characteristics in babbling", *Phonetica*, n°54, p.172-186, 1997.

[16] Rossato, S., Badin, P., Bouaouni, F., "Velar movements in French: an articulatory and acoustical analysis of coarticulation", *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 2003.

[17] Smith, B. L., Brown-sweeney, S. & Stoel-Gammon, C., "A quantitative analysis of reduplicated and variegated babbling", *First Language*, n° 9, p. 175-190, 1989.

[18] Stefanuto, M., Vallée, N., Consonant systems: From universal trends to ontogenesis, *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco, 1999.

[19] Tubach, J.P. & Boë, L.J., *Un corpus de transcription phonétique (300.000 phones) : constitution et exploitation statistique*, Ecole Nationale Supérieure des Télécommunications, Paris, 1990.

[20] Vallée, N., *Systèmes vocaliques: de la typologie aux prédictions,* Thèse de Doctorat, Grenoble, 1994.

[21] Verin, F., *Etudes statistiques de données longitudinales du babillage: cyclicités, co-occurrences, closions*, Mémoire de Maîtrise, ICP, Université Stendhal, Grenoble III, 2001.

[22] Vilain, A., Abry, C., Badin, P., Brosda, S. (1999), 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.