LINGUAL CO-OCCURRENCE CONSTRAINTS IN BABBLING: AN ACOUSTICAL STUDY

Christine L. Matyear, Ph. D.

The University of Texas at Austin
matyear@austin.utexas.edu

ABSTRACT

Acoustical measurements of F2 transitions of 370 babbled CV sequences showed that places of articulation for consonantal closure correlate highly with vocalic tongue frontness/backness. This finding is interpreted as a confirmation of the MacNeilage and Davis Frame/Content theory of speech development.

Keywords: babbling, acoustics, co-occurrnce constraints, Frame/Content theory

1. INTRODUCTION

Babbling, the speech-like, but non-linguistic strings of consonants and vowels uttered by infants between the ages of seven months and approximately eighteen months have been recorded, transcribed, and analyzed by scholars worldwide, and these studies have shown that babbling is phonetically very consistent among infants regardless of the language ambience into which they are born. A production-based theoretical explanation of babbling has been proposed by MacNeilage and Davis, the Frame/Content theory [13], which argues that babbling represents a regularization of infant vocalization by rhythmic close/open movements of the mandible (jaw), a step toward the production of the speech sounds, the consonants and vowels (Cs & Vs) of language. During babbling, they have claimed, infants use this mandibular oscillation almost entirely, to the exclusion of most independent movements of other speech articulators, to produce the percept of consonants and vowels organized into syllable-like CV sequences.

Some of the most striking consequences of this limited production method are the distinctive CV patterns of babbling, e.g. [ba ba ba] or [dæ dæ dæ]. These patterns are characterized by the co-occurrence of labial consonants (consonants made with the lips) with tongue-central vowels, e.g. [ba], coronal (tongue-front) consonants with tongue-front vowels, e.g. [dæ], and dorsal (tongue-back) consonants with tongue-back vowels, e.g. [gu]. These dominant co-occurrences in babbling occur, according to MacNeilage and Davis, because the tongue tends to stay relatively stationary while the jaw makes the movements that produce the closure for consonantal sounds and the opening for vocalic sounds. It is not difficult to understand how these sound sequences are produced.

To MacNeilage and Davis these strong co-occurrence patterns are evidence of motor constraints against the production, during babbling, of individual consonants and vowels in the multitudinous combinations typical of adult speech. They are also evidence of the supremacy of the syllable as the fundamental structural unit of speech. In the process of speech acquisition, they claim, the frame (the syllable) emerges first and only later does the content (the segment). That is, only later does the infant develop the motor control to move the various articulators rapidly and in concert to produce individual consonants and vowels with unique and contrasting articulatory features. These evident production constraints have been reported in transcription studies of babbling infants from language environments as different as English, Spanish, French, Korean, and Japanese.

Transcription studies of the babbling phase of infant speech acquisition have consistently reported similar findings with respect to the repertoire of consonantal and vocalic sounds produced and the characteristic organization of these sounds into CV syllables. The most frequent consonants are the labial and coronal (oral) stops [b] and [d], and the labial and coronal nasal (stop)s [m] and [n] [7, 11, 15, 16, 17, 18]. Vocalic sounds are most often reported to be non-high, non-back [1, 2, 3, 5, 6, 7, 8, 9, 16] As mentioned above, these frequently-occurring consonantal and vocalic sounds also show a strong tendency to co-occur in patterns: labial consonants co-occurring with central vowels, coronal consonants with front vowels, and (though relatively less frequent) dorsal consonants with back vowels [5].

Questions may arise, however, about some of the reported babbling data. These include questions about the validity of using linguistically-based transcription methods for non-linguistic productions, questions about inter-transcriber agreement. The symbols of the International Phonetic Alphabet (IPA) have been developed to permit the phonetic transcription of any speech sound produced in the world's languages, but we might ask whether they are adequate to capture the characteristics of babbling. The degree to which the transcribers in the MacNeilage and Davis study agreed among themselves (personal communication) provides one indication. While agreements about labial sounds [b], were quite solid, more dissention occurred regarding the representation of lingual sounds: coronal [d] and dorsal [g]. In addition, transcribers who did not share the same regional dialect sometimes disagreed about the frontness or backness of vowel sounds. It was generally perceived, however, that the co-occurrence of velar consonants with back vowels were relatively less frequent in the infant corpus overall that the other two forward pairings—labial-central and coronal-front. Still, relying solely on the transcription of infant babbling for fine-grained distinctions in articulation may be unrealistic [4]. This paper addresses some of the issues of inter-transcriber agreement in the transcription of babbling.

Acoustic analysis of vowel formant frequencies can provide additional evidence about tongue position during CV production. The second formant frequency, F2, produced in the vocal tract during the articulation of a vowel is an indicator of tongue position in the horizontal plane (front-central-back). As a result, an acoustical study of babbled CV sequences, measuring the amount of F2 transition (frequency change) between the release of consonantal closure and the full amplitude of vocalic opening, can be revealing. A stable transition for the second formant would imply a lack of tongue movement, regardless of where in the oral cavity the tongue was, and thus the holistic nature of the CV sequence (frame dominance) could be confirmed independently of the transcription. This paper advances the following hypothesis: the F2 transition between the closed and opened phases of lingual CV sequences in babbling should remain close to the same frequency. That is, consonant-vowel (CV) transitions should be relatively flat, indicating an absence of tongue movement in the horizontal plane during the entire CV production.

2. METHOD

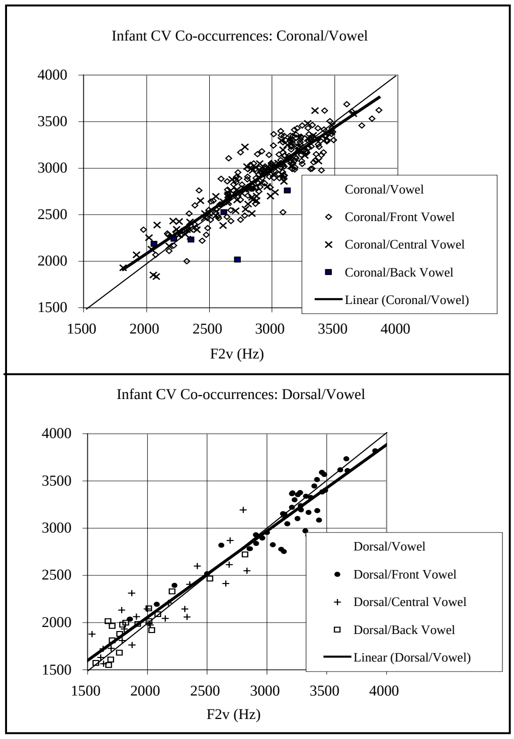
Data were collected from four infants, two female and two male, one-hour recordings made once a week beginning at age seven months and continuing until eighteen months of age. All four were children of mono-lingual American-English speaking parents. Consonant and vowel inventories and CV co-occurrences were noted. A convenience sample of 370 lingual CV(C) sequences was selected for acoustic analysis. These tokens were of the form [dV(d)] or [gV(g)], with front vowels, central vowels, or back vowels in second position. Following the method used by Lindblom [10], measurements were made of F2 vowel onset ($F2_o$) and F2 vowel steady-state ($F2_v$) frequencies. $F2_o$ was measured at the first glottal pulse after the burst, and $F2_v$ at steady state of the vowel. A correlation analysis was made of $F2_o$ as a function of $F2_v$, and a calculation of $F2_v$ - $F2_o$ was performed.

A comparison study was performed with two male and one female adult speakers of American-English. 144 tokens of the same forms as those selected for the infant study were elicited from these subjects. The data were collected using the same recording apparatus, but the tokens were recorded in a laboratory setting. CVC sequences of the desired type were elicited by having the subjects read a prepared list of words containing the appropriate sound sequences and then a narrative passage with the same words included. The same data analyses were performed on these tokens.

3. RESULTS

The frequencies of vocalic $F2_o$ (vowel) onsets and $F2_v$ vowel (steady-states) were measured for the 370 infant and 144 adult CV sequences. As mentioned above, the infant utterances were spontaneous productions transcribed by trained transcribers whose agreement using a point-to-point agreement rating was between 33 - 83% (later, a more global agreement analysis produced much higher ratings: 85 - 98%). The adult utterances were elicited. As Figure 1 shows, a regression analysis of 280 babbles transcribed as [dV(d)] yielded the following result: R^2 = 0.824, SE = 160, y = 0.91x + 262. The analysis of 90 [gV(g)] babbles yielded R^2 = 0.941, SE = 160, y = 97x + 234.

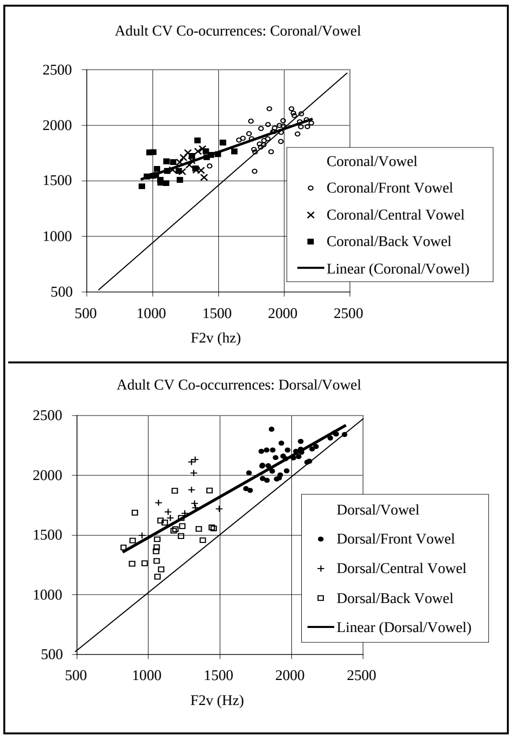
Figure 1: Infant regression analysis.



The striking finding is that both sets of infant data have practically identical slopes, both effectively 1.000, and little variance. These results clearly indicate that for the majority of the utterances measured, the tongue does not move from its position at vowel onset to its position at vowel steady-state. In other words, for most of these utterances the tongue does not move appreciably while the jaw does most of the work in articulating consonant-vowel sequences. For most of these utterances tongue-front consonants do co-occur with tongue-front vowels and tongue-back consonants co-occur with tongue-back vowels even if the transcription does not always capture the fine details of the articulations.

In contrast, Figure 2 shows a regression analysis of 72 adult syllables transcribed as [dV(d)] that yielded the following result: R^2 = 0.744, SE = 94, y = 0.4163x = 1136.6. The analysis of 72 [gV(g)] syllables yielded R^2 = 0.777, SE = 1160, y = 0.681x + 797.49. These adult results show very different characteristics. The slopes of the two data sets are quite different from one another, neither is close to 1.00, and the data points are not as tightly clustered, especially for the [g] tokens. These findings imply a distinctive difference between the articulation of the alveolar consonant and the velar consonant as well as an evident movement of the tongue in the horizontal dimension in order to produce tongue-front consonants with tongue-back vowels and vice versa.

Figure 2: Adult regression analysis.



Finally, a simple comparison of the F2 transition ($F2_v$ - $F2_o$) means for infant vowels and adult vowels shows that infant $F2_v$ values differed from $F2_o$ values by an average of 2 Hz (a non-significant amount), while adult F2 values varied by an average of 281 Hz (a significant amount). It is hard to escape the conclusion that adults evince more independent articulatory control of the tongue and jaw during CV production than infants do.

4. DISCUSSION

Overall the results of these various acoustic analyses leave us with a clearer picture of the infants' developing sound production system. Possibly driven by the same basic motor functions that we use for such ingestive acts as licking, sucking, and chewing, infants at about seven months of age begin to produce speech-like sounds by rhythmically closing and opening the jaws while producing voicing. This powerful dominance of the mandibular cycle permits infants to utter a limited repertoire of CV sequences produced with a common place of articulation, a minimum of articulatory movements, and no evident intentionality.

The infant regression data with their distribution of F2 values along a line with a slope of approximately 1.000 show very convincingly that wherever the tongue is positioned during consonantal closure, it tends to stay there during vocalic opening. It might have been thought that transcribed infant CV syllables that do not show the predicted co-occurrences of Frame/Content theory (e.g. CVs transcribed as [da] or [ge]) would not show a flat F2 transition, and indeed some of them did not. Yet the majority of CV sequences, regardless of their transcribed form, do show relatively flat F2 transitions.

This apparent incompatibility with the predictions of Frame/Content theory demands explanation. It may be the case that listeners have transcribed some tokens too broadly, either because of perceptual ambiguity in the sound or because of the inadequacies of adult transcription conventions coupled with a certain degree of inescapable transcription bias. For example, if the infant's tongue is in a position to touch somewhere on the palate between an alveolar and a velar place of articulation, an English transcriber might very well transcribe an alveolar stop with a central vocalic sound when in fact the infant is producing a post-alveolar stop with a front-central vocalic articulation. It is also possible that an infant's velar articulation without lip rounding might be perceived by an English transcriber as a velar stop with a front or central vowel. Certainly there is room for further investigation.

Regardless, the evidence implies that the infants were producing utterances with little if any independent movement of the tongue. Therefore, though these findings may offer a gentle indictment of the adequacy of adult phonetic symbols for representing infant babbling, they also provide a strong acoustic confirmation of Frame/Content theory co-occurrence constraints: (1) wherever the tongue is for consonantal closure, the tongue tends to remain during vocalic opening, and (2) tongue-front CV co-occurrences are more frequent than tongue-back CV co-occurrences.

5. REFERENCES

[1] Bickley, C. 1983. Acoustic evidence for phonological development of vowels in young children. Paper presented at the Tenth International Congress of Phonetic Sciences, Utrecht, Holland.

[2] Boysson-Bardies, B. de, Halle, P., Sagart, P. L., Durand, C. 1989. A cross-linguistic investigation of vowel formants in babbling. *J. Child Lang*., 16, 1-17.

[3] Buhr, R. D. 1980. The emergence of vowels in an infant. *JSHR*, 231, 73-94.

[4] Cucchiarini, C. 1996. Assessing transcription agreement: methodological aspects. *Clin. Ling. & Phon.*, 102, 131-155.

[5] Davis, B. L., MacNeilage, P. F. 1995. The articulatory bases of babbling. *JSHR*, 38, 1199-1211.

[6] Holmgren, K., Lindblom, B., Aurelius, G., Jalling, B., Zetterström, R. 1986. On the phonetics of infant vocalization. In B. Lindblom, R. Zetterstrom (eds), *Precursors of Early Speech.* New York: Stockton Press, 51-63.

[7] Kent, R. D., Bauer, H. R. 1985. Vocalizations of one-year-olds. *J. Child Lang*., 12, 491-526.

[8] Kent, R. D., Murray, A. D. 1982. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. J. *Acoust. Soc. Am.*, 72, 353-365.

[9] Lieberman, P. 1980. On the development of vowel production in young children. In G. Yeni-Komshian, J. F. Kavanaugh, C. A. Ferguson (eds), *Child Phonology, Vol. I: Production*. New York: Academic Press, 113-142.

[10] Lindblom, B. 1998. An articulatory perspective on the locus equation. *Behav. & Brain Sci.*, 21, 274-275.

[11] Locke, J. 1983. *Phonological Acquisition and Change*. New York: Academic Press.

[12] MacNeilage P. F., Davis, B. L. 1993. Motor explanations of babbling and early speech patterns. In B. Boysson-Bardies, S. de Schoen, P. Juscxyk, P. MacNeilage, J. Morton (eds), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life.* Dordrecht: Kluwer, 341-352.

[13] MacNeilage, P. F., Davis, B. L. 1990a. Acquisition of speech production: frames then content. In M. Jeannerod (ed), *Attention and Performance XIII: Motor Representation and Control. Hillsdale, NJ: Erlbaum*, 453-475.

[14] MacNeilage, P. F., Davis, B. L. 1990b. Acquisition of Speech Production: The Achievement of Segmental Independence. In W. J. Hardcastle, A. Marchal (eds), *Speech Production and Speech Modelling*, Kluwer: Dordrecht, Holland, 55-68.

[15] Oller, D. K. 1986. Metaphonology and infant vocalizations. In B. Lindblom, R. Zetterstrom (eds), *Precursors of Early Speech*, Basingstoke, Hampshire: MacMillan, 21-35.

[16] Roug, L., Landberg, I., Lundberg, L.J. 1989. Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *J. Child Lang.*, 16, 19-40.

[17] Stoel-Gammon, C. 1985. Phonetic inventories, 15-24 months: a longitudinal study, *JSHR*, 28, 505-512.

[18] Vihman, M., Macken, M., Miller, R., Simmons, H., Miller, J. 1985. From babbling to speech: a reassessment of the continuity issue, *Lang.*, 60, 397-445.