Intersyllabic and Word-Level Regularities in Early Acquisition

Peter F. MacNeilage[†] and Barbara L. Davis[‡]

† The University of Texas at Austin, Austin, TX

E-mail: macneilage@mail.utexas.edu, babs@mail.utexas.edu

ABSTRACT

According to the Frame/Content (FC) theory of evolution of speech, phylogeny and ontogeny of speech have always had in common a "Motor Frame" - an elevation-depression cycle of the mandible yielding the mouth close-open alternation manifest in the CV syllable. Infants begin babbling with a phylogenetically old tendency towards *reduplication* of CV cycles and proceed, as in language history, towards the intercyclical *variegation* required in languages. In the process, they develop the syllable as an independent control entity. Infants initially display three biomechanically based vowel-consonant co-occurrence patterns, signifying vowel-consonant interdependence, and thus absence of syllable boundaries. However, languages for the most part do not have these vowel-consonant co-occurrences indicating, according to FC theory that syllable boundaries, and thus syllables, gradually evolved. The ontogeny of syllabification probably recapitulates this phylogeny. Early word-level phenomena, reflected in language patterns, include tendencies for word-initial vowels to be central and for utterance-terminal energy to decrease.

1. INTRODUCTION

The alternation between consonant and vowel is one of the most basic properties of speech. The CV syllable is universal in speech [1] and is the only syllable type in many languages [2]. It is also the most dominant form in speech acquisition, from the beginning of babbling onward. According to the Frame/Content theory of evolution of speech [3] the mandibular elevation/depression cycle that underlies the consonant-vowel alternation constitutes the "Motor Frame" for speech. It may have originated in mandibular ingestive cyclicities in early mammals, circa 200 million years ago, and been used for visuofacial communicative cyclicities (lipsmacks, tonguesmacks and teeth chatters) in ancestral primates, before being paired with phonation to form the protosyllable.

According to this conception, speech may have begun with a Frame stage in which motor frames without much elaboration may have been paired with concepts to form words. We infer a lack of elaboration of the original CV sequences because we have found three patterns of co-occurrence of consonants with following vowels to be ubiquitous in infant babbling and early words and also widespread in languages, which indicates a basic tendency towards lack of non-mandibular articulatory change in CV forms in general. [4] These co-occurrence patterns are; coronals with front vowels, dorsals with back vowels and labials with central vowels. We call the latter pattern "Pure Frames" because it can be produced by mandibular oscillation alone.

As speech evolution progressed, the frame stage may have gradually evolved into the present day Frame/Content stage in which independent segmental elements (consonants and vowels – Content) are programmed into syllable structure frames. This stage is revealed by serial ordering errors in speech, in which misplaced consonants and vowels obey a syllable structure (Frame) constraint on their destinations. The main property of this constraint is that content elements intended for syllable margins (consonants) cannot be misplaced into a syllable nucleus position, and content elements intended for syllable nuclei (vowels) cannot be misplaced into syllable margin positions. The Frame/Content stage presumably evolved from the frame stage in response to selection pressures to increase the number of phonetically distinctive message possibilities well beyond the capacity of frames alone. According to the theory, the infant recapitulates this evolutionary progression in moving from the initial frame stage of babbling to the frame/content mode possessed by adult speakers of her language. We now discuss what this ontogenetic progression involves.

2. FROM CV REDUPLICATION TO VARIEGATION

Reduplication and Variegation in Babbling

It is well known that early babbling tends to be reduplicative; i.e. the infant tends to produce a series of iterations of a single CV form (e.g. "bababa"). It is commonly thought that a predominance of reduplication in early babbling is followed by a predominance of variegation (i.e. successive consonants and/or vowels differ) in late babbling. This is not the case (see [5] for a summary of the evidence). For example, in a study of 6 infants [6] we found that about 50% of the time the syllable following a given syllable was the same in both early and late babbling. We predicted that instances of variegated babbling might result primarily from variability in what we called the "vertical" dimension, involving the attribute of manner (roughly, amount of constriction) for consonants, and the attribute of height for vowels. We reasoned that variability in these attributes might be frame-related; i.e. due only to random variability in elevation and depression cycles of mandibular oscillation, rather than to active changes in position of other articulators. This prediction has been confirmed for both babbling and early speech [4].

ISBN 1-876346-48-5 © 2003 UAB

The Progression Toward Controlled Variegation in Early Words

In a study of the first words of 10 infants, [7] in the period between 12 and 18 months, we found 2 main tendencies toward systematic intercyclical variegation. The first was a tendency to end a word with a high vowel, most notable in the use of diminutive words such as "baby" and "kitty". This tendency may be language specific but may be a result of a greater freedom to produce innovations in absolute final word position because no other articulatory maneuver needs to follow it.

The second tendency involved beginning a word with a labial consonant-vowel-coronal consonant sequence. We also found this tendency in 7 other reports involving 5 different languages [4]. There have even been reports [of infants producing the labial-coronal (LC) sequence when the target word required the opposite sequence [e.g. "pot" for 'top"] [7]. In our study, 9 out of 10 infants preferred LC sequences to CL sequences, with a mean ratio of 2.55:1 [7].

We have argued elsewhere [8] that the LC effect is a self-organization consequence of the simultaneous operation of 3 factors in first word production. First, frames with labial consonants and central vowels may be easier to produce than frames with coronals and front vowels because the latter involves the addition of a tongue fronting movement, while the former only involves frame production. Second, an addition to the functional load associated with vocalization in the babbling stage occurs at the first word stage when an infant has to interface a vocal episode with a lexical concept rather than simply producing an output episode. This produces a generalized bias towards increasing use of pure frames, with their labial-central pattern in first words [8]. Third, the initiation of action is a separable functional role of motor systems, and the execution of this function may become more difficult, the more complex the output pattern is. As a consequence of these three factors, a tendency may arise, in instances in which infants are simulating an adult word with intercyclical variegation, to begin the simulation with the easier pure frame pattern and then add a tongue movement to the next frame.

In order to fully understand the LC pattern, it is necessary to know whether, as in the CV co-occurrence patterns, it is also favored in languages. If it is not, it must be explained in terms of factors specific to the acquisition process, but if it is, it takes on much more significance in the attempt to understand the nature of speech. We studied the occurrence of stops and nasals in CVC words, CVCV words and words beginning with a CVCV sequence in 10 languages (English, Estonian, French, German, Hebrew, Japanese, New Zealand Maori, Quichua, Spanish and Swahili) [7] The LC preference was found in every language except Japanese. The mean ratio of LC to CL patterns was 2.3:1, a ratio almost as high as found infants. The LC preference was also strongly present (8:1) in the putative 27-protoword corpus Bengston and Ruhlen presented [9]. We have suggested that the same factors that may have produced the LC effect in a self-organizational manner in infants may have been at work in earlier hominids, except for the fact that modern infants have a language model to copy, while hominids had to invent it, [5].

The tendency for LC sequences to exceed CL sequences has long been considered to be a part of a "Fronting" phenomenon in word acquisition, first identified by Ingram, according to which the first consonant in a word tends to have a more fronted place of articulation than the second one [10]. The implication of this concept is that there should also be a preference for coronal-vowel-dorsal sequences over their opposite. But we have found this is not a consistent trend, in a study of 11 English-speaking infants [5]. Thus, rather than being describable as fronting, the general pattern in infants is more likely to be a labial-then-lingual one. However, strangely enough, only the predominance of the LC over the CL pattern was present in our sample of 10 languages, with no tendency for LD sequences to exceed DL sequences.

3. THE ORIGIN OF THE SYLLABLE

So far we have described phenomena, which involve more than one CV as intercyclical. But in languages a CVCV sequence is generally considered to involve two *syllables*. However babbling infants do not have syllables if we consider the concept of syllable to include the capacity to place a single syllable type among different syllable types in a syllable sequence. Infants can reiterate particular CV cycles, but as discussed earlier they appear to have virtually no control capacity, which allows them to voluntarily follow one CV cycle with a different one. This ability apparently only begins to develop in the first word stage, initially in the form of the capacity to follow a labial in the first frame cycle with a coronal in the second. However, even at this point, consonant reduplication remains predominant and CV reduplication is frequent.

This lack of intercyclical versatility is indicated by patterns of relationship between adjacent consonants and vowels. As we have described elsewhere in these proceedings [4] there is a strong tendency in infants and languages toward three sets of consonant-vowel co-occurrence patterns that basically reflect a fundamental role of biomechanical inertia in intrasyllabic organization. They are; coronal consonants with front vowels, dorsal consonants with back vowels and labial consonants with central vowels. Babbling and early words tend to be reduplicative. If one now considers the relation between vowels and the consonants that *follow* them in these stages, there would be no biomechanical reason to believe that a vowel would be any less dependent on the consonant that follows it than on the consonant that precedes it. Consequently it would be expected that there would be three sets of vowel-consonant co-occurrence patterns that would be mirror images of the three CV patterns; namely front vowels with coronals, back vowels with dorsals and central vowels with labials. This proved to be the case [11]

However, much to our initial surprise, we found that these 3 vowel-consonant co-occurrence patterns were

absent in our 10-language sample [12]. While the ration of observed-to-expected frequencies of the three CV co-occurrence patterns in our 10 language corpus (where 1.0 is chance) were 1.18, 1.10, and 1.27 for coronal-front, labial-central and dorsal-back respectively, the corresponding figures for VC sequences were 1.04, 1.03 and .88. (The only VC ratios, which were substantially above chance, were front-dorsal, 1.10 and back-labial, 1.11).

We interpret these results to mean that syllable boundaries, between a vowel and the following consonant, and therefore syllables themselves, have evolved in the history of languages as a key step in increasing intercyclical variegation. Such an increase was probably made in response to selection pressures towards an increase in the number of phonetically distinctive message possibilities. These selection pressures have been sufficiently strong to induce generative phonologists to embrace the "Obligatory Contour Principle" (OCP) [13] according to which both transvocalic consonant repetition and transconsonantal vowel repetition are disfavored in modern languages. For example in our study of 10 languages [12], stops and nasals with the same place of articulation were only repeated transvocalically on an average of 67% of the frequency expected by chance. Thus, because of this evolutionary trend in languages, a modern infant is now required to proceed from above chance syllable reduplication characterized by both intracyclical and intercyclical biomechanical constraints, to above chance intersyllabic variegation. They do this as their hominid ancestors did it, by overcoming the inherent biomechanical constraints on VC sequences in the course of developing the syllable as a functionally independent unit.

There may be another reason why intersyllabic variegation has increased to levels beyond those expected by chance. Evidence from psycholinguistics [14] suggests that in modern speech a confusability problem arises in working memory in both perception and production when the same sound recurs at short intervals.

It is worth commenting on the difference in conceptual framework between the present approach to the serial organization of speech and the approach from generative phonology. Generative phonologists observe the ubiquity of intersyllabic variegation in modern languages, give it a name (OCP), which is then considered to have explanatory status, and declare the phenomenon to be innate because it is universal. But there is no consideration of the time domain in generative phonology, and accordingly no principled way of handling the fact that infants begin to speak obeying a principle that is opposite to the obligatory contour principle, and then proceed to turn completely around during development. The frame/content approach reconciles the difference between infants and languages by embracing the basic Neodarwinian tenet of descent by modification. In this view our ancestor's speech constructed a vocal communication system, beginning with an extremely simple starting point (motor frames), and then radically modifying its serial organization by programming the frames with content. The result is embedded in each culture in the form of memes, and subsequently reconstructed by each newly arriving infant using a simulation process. From this standpoint, it is unlikely that the end product of either the deep time (phylogenetic) or the shallow time (ontogenetic) progressions has very much to it that could be called innate.

4. WORD-LEVEL EFFECTS

Words in infants and languages tend to begin with a consonant and end with a vowel. This is not a true word-level phenomenon but a consequence of the existence of a basic Motor Frame consisting of an elevation-depression cycle of the mandible. It is not clear why the cycle involves C followed by V rather than the opposite sequence. Its causes presumably lie more in the realm of production constraints than in perception judging by the tendency of infants with profound hearing deficits to favor the CV form over other alternations [14]. Informal observation of the calls of a number of different mammals suggests that beginning a vocal event with a relatively elevated mandible and then depressing it during most of the phonatory event, and only elevating it again after phonation stops, might be a deep seated pattern.

When words of infants and languages depart from this pattern and begin with a vowel and/or end with a consonant, this seems to have some consequences for the identity of the particular segments concerned. In our study of 10 infants producing their first words [15] the majority of word-initial vowels (64%) were central vowels as compared with 44% of central vowels within utterances. And in our study of babbling of 6 infants [15] 57% of utterance-initial vowels were central as compared with 37% within utterances. In small-scale study of dictionary counts of initial vowels in five languages with 5 vowel systems, we found about 2 1/2 times as many central vowels in word initial position as any of the other 4 vowels [15]. The predominance of central vowels in initial position in babbling may be due to a basic tendency to begin phonation when the tongue is in its resting position in the mouth, before frame generation has begun. Perhaps this was also a basic tendency in earlier hominids, which would account for the prominence of initial central vowels in words of languages (pending further evidence) as well as its prominence in ontogeny.

One well-documented tendency in early final consonants is for them to be more often voiceless than when occurring elsewhere in the utterance. We have found this in a study of single final consonants in babbling [16] and in a study of final consonant clusters in babbling and first words [17]. This tendency is also widespread in languages. We have suggested that this is one of many consequences of a "Terminal Energy Decrease" at the end of an utterance in both babbling and words [16]. In this case a decrease in sub-glottal pressure and/or vocal fold tension could result in a tendency to reduce the pressure drop across the glottis to levels not optimal for voicing. Other common tendencies of infants and languages could be similarly

ISBN 1-876346-48-5 © 2003 UAB

explained [16]. A tendency towards fricatives in final position could result from a terminal reduction in the amplitude of movement of an upper articulator resulting in a tendency for incomplete closure. A tendency towards nasals in final position could result from a tendency to reduce velar elevation resulting in nasal leakage. A terminal drop in fundamental frequency could result, as in a devoicing tendency, from a decrease in subglottal pressure and/or vocal fold tension. A terminal decrease in intensity could result from a decrease in subglottal pressure. In the case of the segmental effects, they must have originated in languages when the words were used in isolation, or in utterance terminal position. Once these patterns were created they remained characteristic of the words even when they occur within running discourse.

5. CONCLUSIONS

According to the Frame/Content theory, certain phenomena are basic to both phylogeny and the ontogeny of speech-like behavior from the start. Most important is the mandibular oscillation underlying the consonant-vowel (CV) alternation (Motor Frame) and its simple reiteration, initially with severe biomechanical constraints against non-mandibular articulatory change. There are some additional initiation and termination propensities – central vowel initiations and a terminal energy decrease. The *syllable* is the main phenomenon that must *subsequently* evolve, and develop in the lifespan of each new speaker, in the service of increase in the number of phonetically distinctive message possibilities

ACKNOWLEDGEMENT

This work was supported in part by National Institutes of Health Grant RO 1 HD2773-09

REFERENCES.

[1] A. Bell and J.B. Hooper, Issues and evidence in syllabic phonology. In A. Bell and J.B. Hooper (eds) *Syllables and segments*, Amsterdam, North-Holland, 1978.

[2] I. Maddieson, *Patterns of Sounds.* Cambridge: Cambridge University Press, 1984.

[3] P.F. MacNeilage, The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21, 499-548, 1998.

[4] B.L. Davis and P.F. MacNeilage, Universal intrasyllabic patterns in early acquisition. (This volume).

[5] P.F. MacNeilage and B.L. Davis in B. Malle, Ed., pp. 118-136, *The Rise of Language out of Pre-language*. Amsterdam, Benjamins, 2002.

[6] B.L. Davis and P.F. MacNeilage, The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38, 1199-1211, 1995.

[7] P.F. MacNeilage, B.L. Davis, C.M. Matyear, and A. Kinney, Origin of speech output complexity in infants and in languages. *Psychological Science* . 10 (5) 459-460, 2000.

[8] P.F. MacNeilage and B.L. Davis, Origin of the internal structure of words. *Science*. 288, 527-531, 2000.

[9] J.D. Bengtson and M. Ruhlen, Global etymologies. In M. Ruhlen, *On the Origin of Language*, Stanford, Stanford University Press, 1994.

[10] D Ingram, Phonological rules in young children. *Journal of Child Language*, 1, 49-64, l974.

[11] B.L. Davis, P.F. MacNeilage P.F. and C.L. Matyear, Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production, *Phonetica*, 59, 75-104, 2002.

12] MacNeilage, B.L. Davis, A. Kinney and C.M. Matyear,. The motor core of speech: A comparison of serial organization patterns in infants and languages. Invited submission to Special Millenium Issue. *Child Development.* 71, 153-163, 2000.

[13] M. Kenstowicz, *Phonology in Generative Grammar*. Oxford, Blackwell, l994.

[14] H.A. McCaffrey, B.L. Davis, P.F. MacNeilage and D von Hapsburg, Multichannel cochlear implantation and the organization of early speech. *The Volta Review*, 101, 5-28.

[15] P.F. MacNeilage and B.L. Davis, Evolution of the form of spoken words. *Evolution of Communication*, 3, 3-20, l999.

[16] M.L. Redford, P.F. MacNeilage and B.L. Davis, Production constraints on utterance-final consonant characteristics in babbling. *Phonetica*, 54, 172-186, 1966.

[17] K.J. Jakielski, B.L. Davis, and P.F. MacNeilage, Acquisition of consonant clusters in babbling and speech, submitted, 2003.