Child Development, January/February 2000, Volume 71, Number 1, Pages 153–163

# The Motor Core of Speech: A Comparison of Serial Organization Patterns in Infants and Languages

*Peter F. MacNeilage, Barbara L. Davis, Ashlynn Kinney, and Christine L. Matyear*

Comparison of serial organization of infant babbling and early speech with that of 10 languages reveals four movement-related design features reflecting a deep evolutionary heritage: (1) the cyclical consonant–vowel alternation underlying the syllable, a "Frame" for speech consisting of mandibular oscillation, possibly evolving from ingestive cyclicities (e.g., chewing) via visuofacial communicative cyclicities (e.g., lipsmacks); (2) three intracyclical consonant–vowel co-occurrence preferences reflecting basic biomechanical constraints—coronal consonants–front vowels, dorsal consonants–back vowels, and labial consonants–central vowels; (3) a developmental progression from above-chance to below-chance levels of intercyclical consonant repetition; (4) an ease-related labial consonant–vowel–coronal consonant sequence preference for word initiation. These design features presumably result from self-organizational responses to selection pressures, primarily determined by motor factors. No explanation for these design features is available from Universal Grammar, and, except for feature 3, perceptual-motor learning seems to have only a limited causal role in acquisition of any design feature.

## INTRODUCTION

The neo-Darwinian theory of evolution by natural selection is generally recognized as providing the only presently viable scientific explanation of the evolution of complex design (Dawkins, 1986). Pinker and Bloom (1990) have recently emphasized that the theory is, in principle, as applicable to language as it is to any other complex design feature of life forms. Can the theory be applied to one of the more accessible facets of language, the design of speech sound systems?

When the question of the design of speech is considered in developmental science, the nature versus nurture issue immediately arises. In the first half of the century, the dominant behavioristic view (Watson, 1925) endorsed nurture as the source of complex behavior, including speech. Olmstead (1971) and Mowrer (1960) considered perceptual-motor learning to be the dominant developmental process of speech. In the last half of the century, the prevailing view in linguistic science entails dominance of both phylogeny and ontogeny of language by nature, in the form of an innate genetically specified Universal Grammar (Chomsky, 1985) including a speech-related generative phonological component consisting of innate distinctive features (e.g., Halle, 1990; Kenstowicz, 1994). An alternative paradigm to these two extremes has recently arisen according to which ". . . biological structure emerges anew within each individual's development from constrained dynamic interactions between genes and various levels of environment, and is not easily reducible to simple genetic and experiential components" (Johnson, 1997, p. 3; see also Oyama, 1985; Thelen & Smith, 1994). According to this perspective, structure tends to emerge on a self-organizational basis, rather than as a simple result of prescription in advance, by either genes or environment.

One novel feature of the present approach is its focus on the question of serial organization rather than simply on segmental units (consonants or vowels) or the distinctive features of linguistics. Although the dominant modern approach to the evolution of speech is focused on vowels (Lieberman, 1984), more attention has been given to consonants in speech acquisition (e.g., Locke, 1983; Vihman, 1996). The present approach is inspired by a classic paper by Lashley (1951) who argued that the problem of serial order (how any given sequence of events is organized?) is the most important problem in the understanding of action systems in general. In this paper we present evidence for four major design features of the serial organization of speech arising from a comparison of babbling and early speech with language patterns, and consider these features in the light of the three alternative perspectives outlined above.

## THE BASIC CONSONANT–VOWEL ALTERNATION (FRAME)

Babbling and early speech have a great deal in common (e.g., Vihman, Macken, Miller, Simmons, & Miller, 1985). Both feature a rhythmic alternation between consonants and vowels. The consonant–vowel alternation gives rise to syllables: a single vowel forms the syllable nucleus, and consonants form syllable mar-

 0009-3920/2000/7101-0018

gins. The rhythmic property, associated with mandibular oscillation, is present in adult languages (Kozhevnikov & Chistovich, 1965), though languages vary in the extent to which a simple consonant–vowel alternation pervades the overall sound structure.

A recent "Frame/Content" theory of evolution of speech gives a central role to this consonant–vowel alternation (MacNeilage, 1998). It takes as a point of departure a salient property of serial ordering errors in adult speech. In these errors, many of which involve "movement" of consonants or vowels (e.g., spoonerisms), there is a syllable structure constraint whereby consonants and vowels never occupy each others' positions in the syllable (Shattuck-Hufnagel, 1979). For example, although consonants and vowels can exchange (e.g., well made → Mel Wade; ad hoc → odd hack), vowels and consonants never exchange with each other (as in no → own). It is argued that this constraint has been present since the origin of independent premotor programming of consonants and vowels, because it arises from the fact that consonants and vowels require incompatible mandibular movements. Depression—mouth opening—is required for vowels; elevation—mouth closing—is required for consonants.

The oscillation of the mandible for speech, which underlies this constraint, may have had an original precursor in early mammals (circa 200 million years ago) in the form of mandibular oscillation for ingestive purposes (chewing, sucking, licking). It may then have been exapted for visuofacial communicative purposes (e.g., lipsmacks, see Redican, 1975) in prehuman primates, and finally have become paired with phonation to form protosyllables in hominids. The oscillation of the mandible is regarded as the "Frame" for speech. Both speech phylogeny and ontogeny are regarded as primarily a matter of developing internal "Content" for frames, in the form of relatively independent movement components, which eventually give rise to various consonants and vowels.

In this paper we distinguish between intracyclical or frame-internal phenomena—relations between adjacent phases of the cycle—and intercyclical or multiframe phenomena—cases in which at least one phase of a cycle repeats itself. In considering intracyclical phenomena, we focus on the consonant–vowel (CV) sequences rather than the vowel–consonant (VC) sequence for a number of reasons. First, sequences of alternations between consonants and vowels in babbling and early speech tend to begin with consonants and end with vowels. Consequently a more comprehensive view of the alternation process in infants is obtained by considering CV sequences than VC sequences. Second, words across languages also reflect a preference for beginning with a consonant and ending with a vowel (Bell & Hooper, 1978), suggesting an important commonality between infants and adults that invites investigation. A third, related consideration, is that the CV sequence has typically been considered the most important unit in speech beyond the individual segment (consonant or vowel; Bell and Hooper, 1978), often given the status of the only universal syllable type. Despite the focus on the CV sequence in our analyses, we consider that it is the alternation itself which has fundamental phylogenetic and ontogenetic status, beyond the question of how it may begin or end. This status is reflected in the fact that utterances/words beginning with a vowel and ending with a consonant are not rare in the alternating patterns of infants or languages. Consistent with our primary interest in the alternation as such, we also present some information on VC sequences in infants and adults.

## INTRACYCLICAL ORGANIZATION: CONSONANT–VOWEL CO-OCCURRENCE CONSTRAINTS

We begin with some information about consonants and vowels, the components of cycles. We will be primarily concerned with stop consonants and nasal consonants as they are by far the most common consonants in babbling and early speech, and are common across languages. They therefore seem to have a basic status. Stop consonants and nasals are both produced with a total occlusion of the vocal tract (mouth cavity), but in nasals the air passage to the nose remains open. Our primary concern will be with the place of articulation of these consonants—where in the vocal tract occlusion occurs. For labials, closure is made at the lips. English has three labial stops/nasals, the sounds spelled as "p," "b," and "m." For coronals, closure is made in the front of the mouth cavity, roughly on the hard palate, sounds spelled "t," "d," and "n." For dorsals, closure is made in the region of the soft palate—sounds spelled "k," "g," and "ng."

For vowels we will be primarily concerned with tongue position in the front–back dimension of the mouth. Front vowels include those in the words "beet," "bet," and "bat." Central vowels occur in the words "but" and the first syllable of "father." Back vowels are in the words "boot," "boat," and "pot."

*Babbling and early speech.* Babbling, a universal phenomenon that typically begins at about 7 months of age, is defined in terms of relatively rhythmic cycles of alternation between a closed and open mouth configuration accompanied by phonation (vocal fold vibration; Oller, 1986). The prototypical babbling episode consists of a repeated rhythmic alternation between the same open and closed mouth configurations, as in "babababa." Recent work on the details of single cycles in babbling and early speech at the intracyclical level, has revealed three CV co-occurrence preferences: coronal consonants co-occur with front vowels (e.g., "day"); dorsal consonants with back vowels (e.g., "go"); and labial consonants with central vowels (e.g., "ba"). These co-occurrences have been shown in a series of quantitative case studies of a total of 15 participants, during prespeech babbling and early speech. The database for a single participant always exceeded 1000 syllables (Davis & MacNeilage, 1990, 1994, 1995; MacNeilage & Davis, 1996; Zlatic, MacNeilage, Matyear, & Davis, 1997). There have been 48 individual instances of the three types of co-occurrence preferences, and only five instances of any of the other six co-occurrence possibilities.

Table 1 shows median observed-to-expected ratios for the three types of co-occurrence in prespeech babbling (B) (Davis & MacNeilage, 1995), early words during the so-called "50-word stage" from 12 to 18 months (W), and babbling concurrent with words at this stage (C). Expected frequencies for each cell were computed from the overall frequencies of the particular consonant and vowel in the total corpus. Ratios for the three favored co-occurrence types ranged from 1.18 to 1.84 with a median of 1.26. Note that there were only three instances of above-chance median levels of preference for all of the other six possible CV types combined.

A variety of papers in preparation by ourselves and our colleagues suggests that this pattern is widespread across different language environments and different populations. It has been found in an analysis of 3 groups—5 French, 5 Swedish, and 5 Japanese infants—from the Stanford University database (see Boysson-Bardies et al., 1992) by Davis, MacNeilage, Gildersleeve-Neumann, and Teixeira (1999). It has been found in 1 of 2 infants in a Brazilian Portuguese environment (Teixiera & Davis, 1999) and in a study of 7 infants in an Ecuadorian Quichua environment (Gildersleeve-Neumann & Davis, 1998). It has also been found in an infant with a severe-to-profound hearing loss who had received a multichannel cochlear implant (McCaffrey, Davis, MacNeilage, & von Hapsburg, in press) and in a group of 4 infants with extreme speech delay (Davis & MacNeilage, 1999).

**Table 1 Median Ratios of Observed-to-Expected Frequencies of Consonant–Vowel Co-occurrence Types in Prespeech Babbling, First Words, and Babbling Concurrent with First Words**

| | | Consonants | | |
|---|---|---|---|---|
| Vowels | | Coronal | Labial | Dorsal |
| Front | B | 1.28 | .57 | .95 |
| | C | 1.21 | .85 | .89 |
| | W | 1.18 | .75 | .66 |
| Central | B | .84 | 1.34 | .96 |
| | C | .86 | 1.27 | .89 |
| | W | .85 | 1.20 | 1.10 |
| Back | B | .64 | 1.22 | 1.22 |
| | C | .85 | .79 | 1.84 |
| | W | 1.08 | .76 | 1.24 |

*Note:* B = prespeech babbling, C = babbling concurrent with first words, W = first words.

A number of other studies of CV co-occurrences in babbling and early speech by other investigators have produced many confirmations of our findings but also counterexamples and null findings (Boysson-Bardies, 1993; Oller & Steffans, 1993; Tyler & Langsdale, 1996; Vihman, 1992). These studies have uniformly involved much smaller databases per infant than in the above studies, sometimes used different vowel classifications (Tyler & Langsdale, 1996; Vihman, 1992), and sometimes did not take the overall frequencies of both the consonantal and vowel categories into account when computing expected frequencies for individual CV classes (Oller & Steffans, 1993; Tyler & Langsdale, 1996). We believe that if our studies are replicated in terms of database size, vowel classification conventions, and analysis procedures, these three co-occurrence patterns will typically be observed.

As was expected from our conception of the frame in infants as a relatively simple undifferentiated oscillation of the mandible, VC co-occurrence patterns were found to be similar to CV patterns in the Davis and MacNeilage (1995) corpus. In an analysis of 5,573 sequences, the observed-to-expected ratios of the three co-occurrence patterns were: coronal–front vowel, 1.34; dorsal–back vowel, 1.10; and labial–central vowel, 1.36 (unpublished observations).

*Languages.* Janson (1986) studied consonant–vowel relationships derived from written texts of five languages: Finnish, Turkish, Latin, Latvian, and Setswana. Maddieson and Precoda (1992) studied consonant–vowel relationships derived from dictionary counts in five additional languages: Hawaiian, Rotokas, Piraha, Kadazan, and Shipibo. We analyzed the combined data from these two studies, which constituted a total of 205 CV groupings (MacNeilage & Davis, 1993). There was a significant tendency for dentals/alveolars (coronals) to favor front vowels and disfavor back vowels, and a significant tendency for velars (dorsals) to disfavor front vowels. There was also a nonsignificant trend for dorsals to

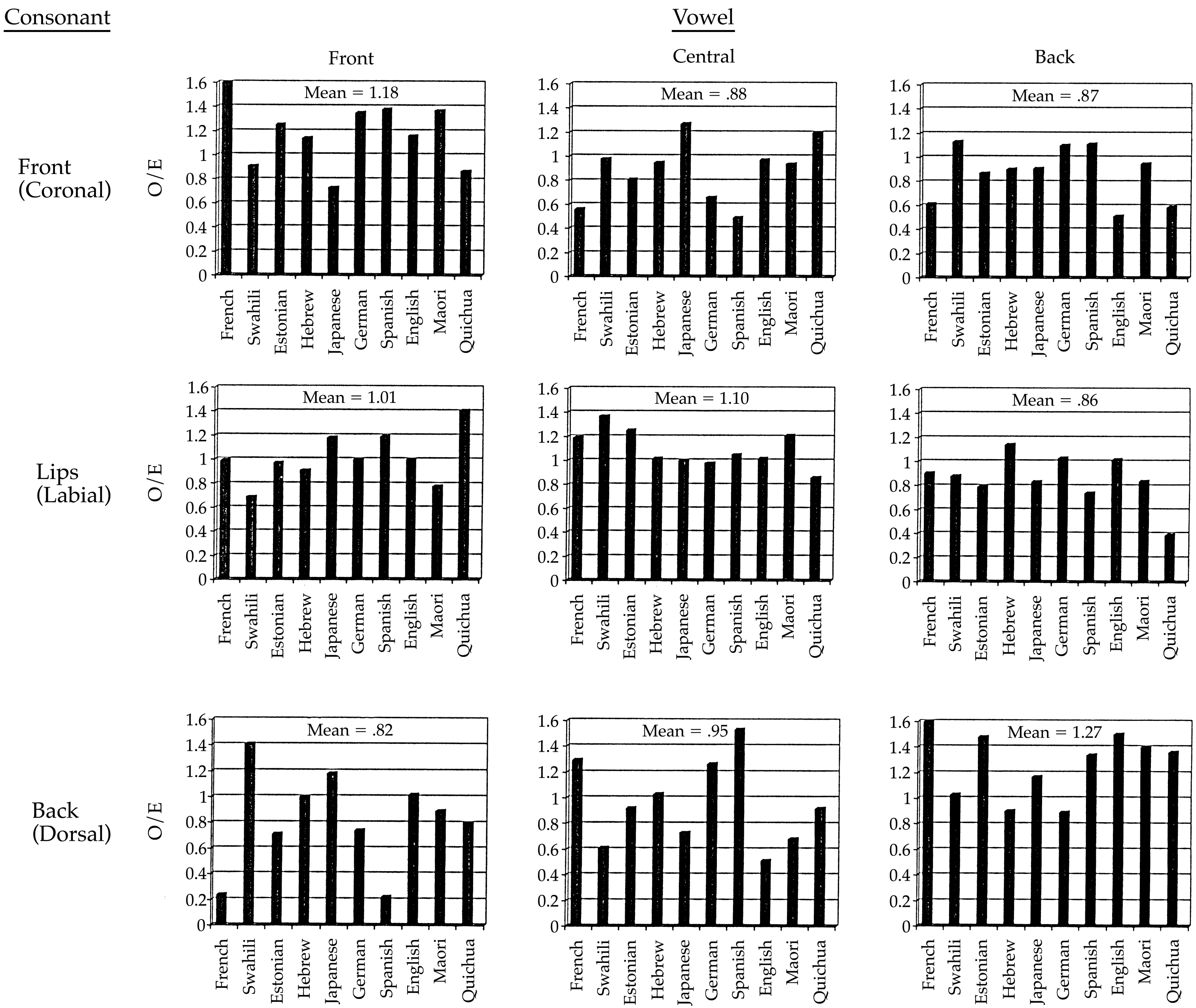


**Figure 1** **Observed-to-expected ratios of 9 consonant–vowel (CV) co-occurrence types in 10 languages.**

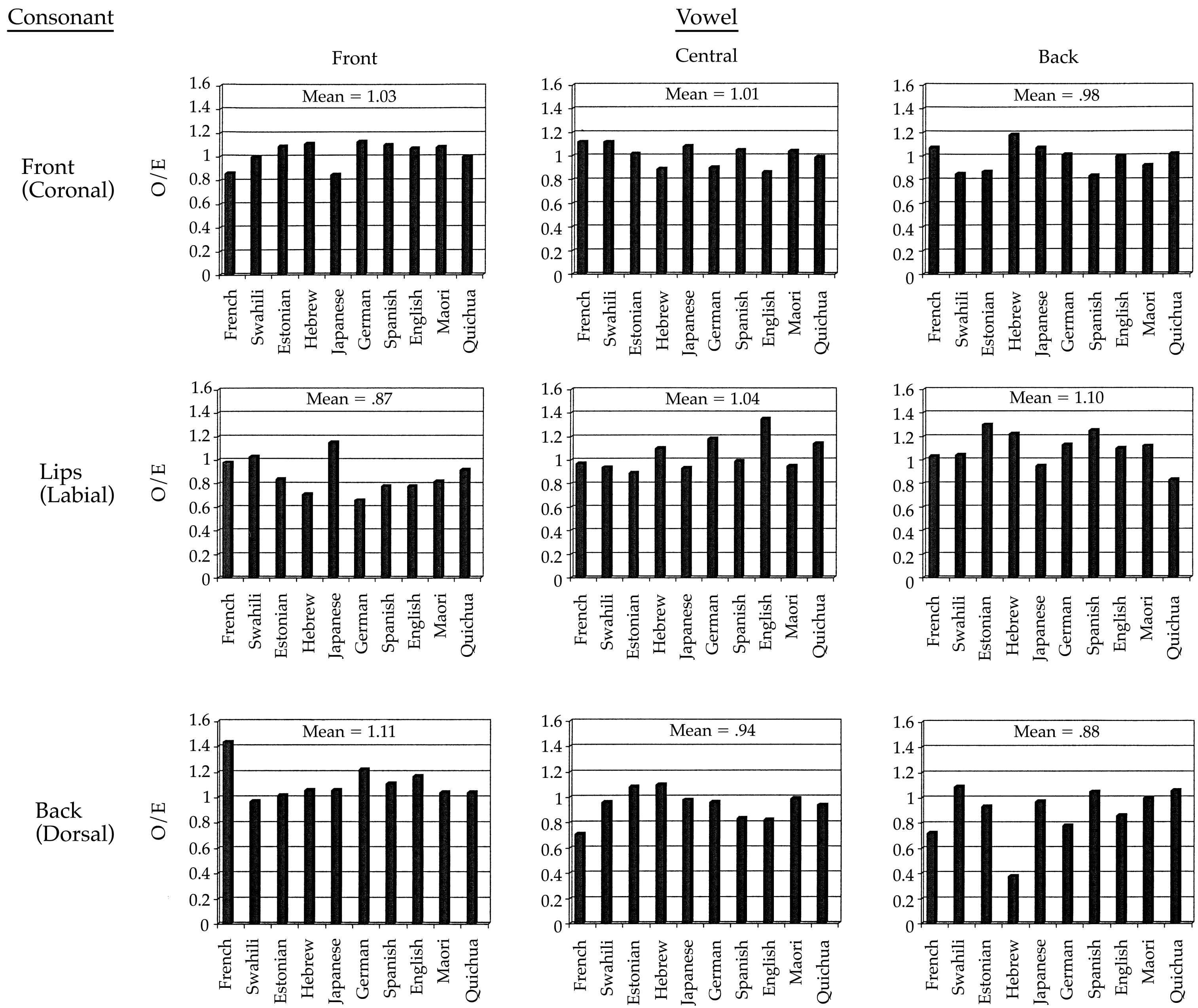


**Figure 2 Observed-to-expected ratios of 9 vowel–consonant (VC) co-occurrence types in 10 languages.**

favor back vowels. In contrast with the trends toward the lingual co-occurrence constraints, there was no obvious sign of the infant tendency for labial consonants to occur with the central vowel /a/.

We have also analyzed intrasyllabic trends involving stop consonants, nasals, and vowels in 10 languages other than the ones studied by Janson (1986) and Maddieson and Precoda (1992): English, Estonian, French, German, Hebrew, Japanese, New Zealand Maori, Quichua, Spanish, and Swahili. For English, the Oxford Psycholinguistic Database, derived from the Shorter Oxford English Dictionary, was used. Dictionaries were the source of data for other languages. A list of the sources used, and of the stops, nasals, and vowels in each language is available at http://homepage. psy.utexas.edu/faculty/macneilage/langs.html. We analyzed CVC words and initial CVC sequences in CVCV words and words that began with a CVCV sequence. A total of 12,630 words was analyzed. The inventory sizes were English, 2,348; Estonian, 477; French, 2,528; German, 1,223; Hebrew, 248; Japanese, 284; New Zealand Maori, 2,481; Quichua, 971; Spanish, 1,242; and Swahili, 828. As the findings from the CVC, CVCV, and CVCV . . . patterns in English were similar, the results from the three word types were pooled in analysis of each language.

The ratios of observed-to-expected frequencies of word–initial CV sequences of labial, coronal, and dorsal consonants with front, central, and back vowels in the 10 languages are shown in Figure 1. The only CV types in which the observed frequencies for the majority of languages exceeded the expected frequencies were the favored CV types in babbling and early speech: labials–central vowels, coronals–front vowels, and dorsals–back vowels. Observed frequencies exceeded expected frequencies in labial–central vowel pairs in seven languages, coronal–front vowel pairs in seven languages, and dorsal–back vowel pairs in eight languages. Although these three categories had 22 instances of above-chance frequencies out of 30 possibilities, the other six categories only had 17 instances of above-chance frequencies out of 60 possibilities. The overall distribution of occurrences was significantly beyond chance, $\chi^2(1, N = 12{,}630) = 13.49$, $p < .01$. Mean values were coronal–front, 1.18; labial–central, 1.10; dorsal–back, 1.27. Every language except Japanese had an overall average that was above 1.0 for the three categories combined.

A very different pattern was observed for VC co-occurrences. As can be seen from Figure 2, none of the three VC patterns favored by infants was clearly favored by the group of 10 languages.

*Explanations.* First, consider why these particular consonants and vowels tend to co-occur in babbling and early speech. Two of the three patterns we observed, coronal consonants with front vowels and dorsal consonants with back vowels, involve the tongue for both the consonants and the vowels. We have called these two types of lingual frames "Fronted Frames" and "Backed Frames" respectively, and have suggested that they may be achieved primarily by placing the tongue in the front or back of the mouth, perhaps typically *before the utterance even begins* (MacNeilage & Davis, 1990). Consequently, the tongue may play only a negligible active role in these patterns during the actual utterance.

The reason for the two lingual co-occurrences seems straightforward from a biomechanical standpoint. The tongue is simply placed in a nonresting position in the front–back dimension and not actively moved during a syllable. The third frame pattern found in infants is the tendency for labial consonants to co-occur with central vowels. Achievement of closure or constriction for labial consonants, which occurs at the lips, has no consequences for tongue position in the front–back dimension. We have argued that the pattern of labial consonants and central vowels involves "Pure Frames." That is, the pattern may be produced by mandibular oscillation alone, without any active tongue movement. The tongue may simply occupy a resting position in the center of the mouth (MacNeilage & Davis, 1990). Such a movement pattern seems very likely to have been part of the earliest speech of hominids, for it can be considered the most basic movement pattern of the most basic oral articulator.

We believe that the occurrence of these three CV patterns found in languages as well as in infant babbling and early speech is of fundamental importance because it probably means that the patterns have been present since the origin of speech in hominids. The presence of the two lingual patterns in languages as well as in infants suggests that a constraint against rapid changes in tongue position has been present since speech production first began.

The presence of VC co-occurrence constraints similar to the CV constraints in infants but not in adults also seems to have important implications. It suggests that the tendency of the syllable boundary to occur after the vowel rather than before in adult language is the result of an ontogenetic progression. To the extent that early speech patterns of infants are like early speech patterns of hominids, syllable boundaries, and therefore syllables themselves, may have also been subject to a phylogenetic progression.

An alternative treatment of a number of phenomena discussed here, namely the frame and the CV and

VC co-occurrence constraints, could be provided by generative phonology, for example, in terms of the most recently favored approach known as Optimality Theory (e.g., Archangeli & Langendoen, 1997). Such treatments either explicitly or implicitly involve the Chomskyan assumption of innate genetically specified mental propensities of humans, often called "innate knowledge" (e.g. Archangeli & Langendoen). For instance, the syllable has sometimes been considered innate, and the CV and VC co-occurrences could also be considered innate, and conceptualized in terms of "Markedness." Markedness is typically used as a summary term for certain sounds and sound patterns occurring more frequently than others (Kenstowicz, 1994). This approach, however, does not result in an explanation in the ordinary sense of the term. As George Miller (1990, p. 321) has pointed out, "linguists tend to accept simplifications as explanations," rather than using the term "explanation" to denote causes of a phenomenon. For example, sound preferences are described in terms of markedness, but then, in a further step, markedness is regarded as an explanation for the phenomenon that it was chosen to denote, with no attention to the possible causes of the phenomenon. A further problem in the generative phonology approach is that the designation of "innate" is considered to apply to aspects of speech regardless of the age of the speaker, and this conception is not compatible with the presence of differences between infant and adult forms, each of which may be universal, such as was observed here for VC co-occurrences.

Another possible alternative approach to these findings is one based on perceptual-motor learning. One fundamental problem with a perceptual-motor learning approach is that it does not deal with the question of how the speech forms that must be learned by the infant evolved in the first place. Nevertheless, with respect to acquisition, it could be argued that infants may have learned the closed–open alternation pattern of adult speech and the CV co-occurrence constraints. The late onset and abnormal form of babbling in hearing-impaired infants (Oller & Eilers, 1988), together with reports that some profoundly hearing-impaired infants do not babble at all (Oller, Eilers, Bull, & Carney, 1985; Osberger et al., 1991), shows that auditory experience does play a key role in babbling. But the uniform dominance of a simple closed–open alternation in infant babbling and early speech, in the presence of great cross-language variation in the complexity of the syllable structure of the ambient language, suggests that if infants indeed learn these patterns by 7 months of age, they have a remarkable capacity to perceptually derive the simple cycle from the input, regardless of its complexity. In addition, it would be necessary to explain why infants show the same co-occurrence for CV and VC sequences when languages only show these patterns in CV sequences. A further problem is the presence of two co-occurrence patterns in the Japanese infants (Davis et al., 1999), which were not in the language (see Figure 1).

## DEVELOPMENTAL DECREASE IN INTERCYCLICAL CONSONANT REPETITION

*Babbling and early speech.* At the intersyllabic level infants have a strong tendency toward "reduplication," a tendency to repeat consonant–vowel sequences (e.g., "bababa"). Davis and MacNeilage (1995) analyzed every available two-syllable sequence produced by 6 participants during the babbling stage. The median level of syllable repetition was 50%. The median level of consonant repetition was about 67%. These levels were well above chance expectations based on the overall frequencies of the particular consonants and vowels being repeated. MacNeilage, Davis, and Matyear (1997) found that the level of consonant repetition in a study of 4 of these participants during the 50-word stage remained about the same as in babbling. The resultant well-known tendency of infants to repeat the same place of articulation of consonants in first words even when the target word has a sequence of dissimilar consonants has been termed "consonant harmony" (e.g., "duck" → "guck," Vihman, 1996).

*Languages.* In contrast with infant preferences, it is well known that transvocalic consonant repetition tends to be disfavored in adult languages (Vihman, 1978). This fact is delineated in the Obligatory Contour Principle of phonology according to which successive instances of the same entry in the consonant tier are banned (Kenstowicz, 1994). A study of the first and second stops and nasals in CVC, CVCV, and CVCV . . . words of our 10 language corpus showed that in only 1 of 30 instances (10 languages, 3 places of articulation) did the tendency to repeat a consonant with the same place of articulation exceed the chance value based on the relative frequencies of the consonant type concerned in the overall corpus. The average tendency over the 30 instances was only 67% of chance values, ranging from .44 in Hebrew to .89 in Swahili. All ratios were significantly below chance levels ($p < .05$).

*Explanations.* Intuitively, one might have expected that as infants grew into adults they might reduce consonant harmony to chance levels as part of a progression toward the levels of serial output complexity necessary for the language to generate a large message set. But why would these levels be reduced to *below-chance* values in adults? If the first hominid speech was strongly reduplicative, one might expect that there would be some residue of this preference in modern languages. This suggests that a problem might arise in modern high-speed speech reception and production that was not present when speech was produced at lower speeds with smaller inventories. The problem may lie in the confusing effect of frequent recurrence of the same sound in working memory, probably in both the stage of input analysis and that of output organization. A classic finding in working memory studies is the confusibility of simultaneously held items with similar pronunciation (Conrad & Hull, 1964). With respect to output, studies of speech errors show that they are potentiated by a "repeated phoneme effect" (MacKay, 1987): the occurrence of two examples of the same sound in close proximity tends to induce serial ordering errors.

A generative phonological approach would seem once again to have a problem explaining the difference between these infant and adult patterns. The Obligatory Contour Principle was formulated by observations of adult speech, and following the procedure noted by Miller it has been used to explain adult speech patterns. But it is incompatible with infant speech, for which the opposite principle is required. The perceptual-motor learning approach also appears to have a problem accounting for these results. A successful perceptual-motor simulation of adult patterns would result in below-chance levels of intercyclical repetition, not above-chance levels.

## THE LABIAL CONSONANT–VOWEL–CORONAL CONSONANT SEQUENCE PREFERENCE

*Infant speech.* Ingram (1974) described the most well-known procedure whereby infants begin to escape from consonant harmony in early words as "Fronting." When consonants are not repeated, the first consonant tends to have a more anterior (front) place of articulation than the second. As dorsal consonants tend to remain relatively rare in early words, the main manifestation of fronting is a labial consonant–vowel–coronal consonant (LC) sequence. In a review of 7 reports involving 5 different language communities (MacNeilage & Davis, in press) the LC tendency was observed in 21 out of 22 infants. Two infants even produced adult words that had the opposite sequence with the LC sequence (e.g. "top" produced as "pot"; Jaeger, 1997; Macken, 1978). We found that 9 of 10 infants in the first 50-word stage also showed the LC tendency (MacNeilage, Davis, Kinney, & Matyear, 1999).

*Languages.* Locke (1983) has considered the relative role of labial and coronal consonants in adult speech sequences. He found strong signs of an LC effect in English. LC sequences were much more frequent (68) than CL sequences (28). The effect was also present in French.

Ratios of LC to CL sequences in the 10 languages we have studied have been reported elsewhere (MacNeilage et al., 1999). They ranged from a high value of 3.33 for Quichua to a low value of .84 for Japanese, the only one of the 10 languages to show a below-chance ratio. Chi-square tests showed that 8 of the remaining 9 languages had significantly more LC sequences than CL sequences. The exception was Swahili which had a nonsignificant 1.34 ratio. The value for English was 2.55 which is similar to Locke's (1983) value of 2.43. The mean value for the set of 10 languages was 2.18.

*Explanations.* Why does this pattern occur in infants? Some considerations suggest that a labial onset is simpler to produce than a coronal (or dorsal) onset. According to the conception of frames presented earlier, labials may be made simply with one phase of a cycle of mandibular oscillation, whereas coronals require an additional tongue movement. There is also evidence of a higher frequency of labials than coronals in first words. Boysson-Bardies et al. (1992) found this trend in groups of 5 infants from each of 4 language communities, and we found it in 3 of 4 infants during the first 50-word period (MacNeilage et al., 1997). This trend can be interpreted as a regression toward easier production forms in the face of the new functional demand to interface the motor system with the mental lexicon. An analogous example of apparent simplification of operation at the signal processing level in the presence of demands associated with concurrent building of a mental lexicon has recently been reported for speech perception by Steger and Werker (1997). In addition, three studies of infants prevented from vocalization during the babbling and early speech periods by early tracheostomies report a very strong preference for labials over coronals in their initial post-tracheostomy speech efforts (Bleile, Stark, & Silverman-McGowan, 1993; Locke & Pearson, 1990; Vaivre-Douret, Le Normand, & Wood, 1995).

That LC preference is also prominent in adult speech forms, even though they typically produce a word immediately after another word rather than beginning a word from silence, as young infants typically do, suggests that the pattern is of fundamental importance to speech. We have suggested that the LC sequence is a self-organizational response to pressures toward increased serial output complexity of speech in the service of increasing the size of the communicable message set (MacNeilage et al., 1999). If it is indeed easier to produce than the CL sequence, it may have been more likely to have occurred adventitiously in early stages of evolution, thus potentiating its adoption for lexical purposes. In infants it may initially be induced nonspecifically as a self-organizational response to the serial complexity of the ambient language, as an infant develops more versatility in output organization by gaining more independent control of individual articulators during the course of a speech utterance.

Ingram (1974), working in the generative phonological tradition, designated Fronting as a manifestation of markedness. As we have noted, however, this designation does not have explanatory value. The LC sequence also presents an explanatory problem for the perceptual-motor learning approach. This approach cannot explain the presence of the pattern in adult languages. Moreover, the strength of the pattern tends to be greater in infant babbling and early speech than in adult language, perhaps especially so in the 2 infants who reversed the coronal–labial sequences in the adult words they were trying to say (Jaeger, 1997; Macken, 1978). The mean labial–coronal preference ratio in languages was 2.18. The median LC preference ratio in the 10 infants we studied (MacNeilage et al., 1999) was 6.75, and 3 of the infants had *no* words with CL sequences. Perceptually based learning is therefore unlikely to be the only variable involved.

## CONCLUSIONS

In the context of the nature–nurture question, we have considered three possible types of explanation for the form of babbling and early speech, and adult language patterns: genetic determination via Universal Grammar, perceptual-motor learning, and self-organization. We identified four design features of speech organization as central phenomena that need to be explained. We conclude that these speech patterns are best understood in primarily self-organizational terms. In our view, the normal context of speech acquisition results, from the beginning of babbling onward, in the manifestation of motor propensities more basic than can be attributed to speech-specific genetics, and simpler (early frames) or stronger (LC effect) than can be attributed to learning alone. In addition, with both the genetic and the learning approaches there is a problem explaining the fact that some early serial organization patterns are similar to ambient language patterns (CVs, LC effect) but others are not (VCs, reduplication propensities).

With respect to the nature–nurture issue, the evidence suggests that "nurture has nature" (see Plotkin, 1997). The incorporation in languages of these basic patterns, patterns that apparently recur in each generation of infants, gives the infant a ready-made initial access to the ambient language. The propensities, although neither primarily genetically specified for speech nor primarily learned, reduce the infant's initial task to one of fitting specific available output patterns to adult words that have similar patterns.

The evolutionary implications of the work reported here arise primarily from the claim that the common presence of Frames, intracyclical co-occurrences, and LC sequence preferences in infants and languages (features 1, 2, and 4) suggests their presence in the earliest language/s. The extremely basic nature of these design features makes it difficult to argue that they only arose in later stages of speech phylogeny. The developmental decrease in relative frequency of consonant repetition (feature 4) may be a present-day reflection of the fact that evolutionary increases in speech complexity produce problems of serial confusibility. Consequently, modern infants must slowly learn motorically uncongenial sequences which have evolved in the cultural matrix of each language in response to pressures toward reduced serial confusibility.

## ACKNOWLEDGMENT

This work was supported in part by National Institutes of Health grant R01 HD2773.

## ADDRESSES AND AFFILIATIONS

Corresponding author: Peter F. MacNeilage, Department of Psychology, The University of Texas at Austin, Austin, Texas 78712; e-mail: macneilage@mail.utexas.edu. Barbara L. Davis, Ashlynn Kinney, and Christine L. Matyear are also at the University of Texas at Austin.

## REFERENCES

Archangeli, D. B., & Langendoen, T. (Eds.). (1997). *Optimality theory: An overview.* Oxford, UK: Blackwell.

Bell, A., & Hooper, J. B. (Eds). (1978). *Syllables and segments.* Amsterdam: North-Holland.

Bleile, K. M., Stark, R. E., & Silverman-McGowan, J. (1993). Speech development in a child after decanulation: Further evidence that babbling facilitates later speech development. *Clinical Linguistics and Phonetics, 7,* 319–330.

Boysson-Bardies, B. (1993). Ontogeny of language-specific syllable production. In B. Boysson-Bardies, S. de Shonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Boysson-Bardies, B., Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I., & Arao, F. (1992). Material evidence of infant selection from the target language: A cross-linguistic study. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications.* Timonium, MD: York Press.

Chomsky, N. (1985). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology, 55*, 429–432.

Davis, B. L., & MacNeilage, P. F. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research, 33*, 16–27.

Davis, B. L., & MacNeilage, P. F. (1994). Organization of babbling: A case study. *Language and Speech, 37*, 341–355.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 38*, 1199–1211.

Davis, B. L., & MacNeilage, P. F. (1999, June). *Integration of perceptual and motor information in speech acquisition.* Paper presented at the Speech Communication and Language Development Symposium, Stockholm, Sweden.

Davis, B. L., MacNeilage, P. F., Gildersleeve-Neumann, C., & Teixeira, E. (1999, July). *Cross-language studies of consonant-vowel co-occurrence constraints in infants and adults: Ambient language effects in first words.* Paper presented at the 20th Annual Child Phonology Conference, Bangor, Wales.

Dawkins, R. (1986). *The blind watchmaker.* New York: Norton.

Gildersleeve-Neumann, C., & Davis, B. L. (1998). *Production versus ambient language influences on speech development in Quichua.* Paper presented at the Annual Meeting of the American Speech, Hearing, and Language Association, San Antonio, Texas.

Halle, M. (1990). Phonology. In D. N. Osherson & H. Lasnik (Eds.), *Language: An invitation to cognitive science*. Cambridge, MA: MIT Press.

Ingram, D. (1974). Fronting in child phonology. *Journal of Child Language, 1*, 233–241.

Jaeger, J. (1997). How to say 'Grandma': The problem of developing phonological representations. *First Language, 17*, 001–029.

Janson, T. (1986). Cross-linguistic trends in CV sequences. *Phonology Yearbook, 3*, 179–196.

Johnson, M. H. (1997). *Developmental cognitive neuroscience.* Oxford, UK: Blackwell.

Kenstowicz, M. (1994). *Phonology in generative grammar.* Oxford, UK: Blackwell.

Kozhevnikov, V. A., & Chistovich, L. (Eds.). (1965). *Speech: articulation and perception.* Clearing House for Federal Scientific and Technical Information. Joint Publications Research Service, 30:543.

Lashley, K. (1951). The problem of serial order in behavior. In L. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley.

Lieberman, P. (1984). *The biology and evolution of language.* Cambridge, MA: Harvard University Press.

Locke, J. L. (1983). *Phonological acquisition and change.* New York: Academic Press.

Locke, J. L., & Pearson, D. (1990). Linguistic significance of babbling: Evidence from a tracheostomized infant. *Journal of Child Language, 17*, 1–16.

MacKay, D. (1987). *The organization of perception and action.* New York: Springer-Verlag.

Macken, M. (1978). Permitted complexity in phonological development: One child's acquisition of Spanish consonants. *Lingua, 44*, 219–253.

MacNeilage, P. F. (1998). The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499–548.

MacNeilage, P. F., & Davis, B. L. (1990). Acquisition of speech: Frames, then content. In M. Jeannerod (Ed.), *Attention and performance XIll.* Hillsdale, NJ: Erlbaum.

MacNeilage, P. F., & Davis, B. L. (1993). Motor explanations of babbling and early speech patterns. In B. de Boysson-Bardies, S. deSchonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Changes in speech and face processing in infancy: A glimpse at developmental mechanisms of cognition.* Dordrecht: Kluwer Publishing Company.

MacNeilage, P. F., & Davis, B. L. (1996, May). *From babbling to first words: Phonetic patterns.* Paper presented at the first ESCA tutorial and Research Workshop on Speech Production Modelling.

MacNeilage, P. F., & Davis, B. L. (in press). Evolution of speech: The relation between ontogeny and phylogeny. In L. Knight & J. R. Hurford (Eds.). *The evolutionary emergence of language.* Cambridge, UK: Cambridge University Press.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (1999). Origin of serial output complexity in speech. *Psychological Science, 10*, 459–460.

MacNeilage P. F., Davis, B. L., & Matyear, C. L. (1997). Babbling and first words: Phonetic similarities and differences. *Speech Communication, 22*, 269–277.

Maddieson, I., & Precoda, K. (1992). Syllable structure and phonetic models. *Phonology, 9*, 45–60.

McCaffrey, H. L., Davis, B. L., MacNeilage, P. F., & von Hapsburg, D. (in press). Effects of multichannel cochlear implant on the organization of early speech. *The Volta Review.*

Miller, G. A. (1990). Linguists, psychologists and the cognitive sciences. *Language, 66*, 317–322.

Mowrer, O. H. (1960). *Learning theory and symbolic processes.* New York: Wiley.

Oller, D. K. (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech.* Basingstoke, Hampshire, UK: MacMillan.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59*, 441–449.

Oller, D. K., Eilers, R. E., Bull, D. H., & Carney, A. E. (1985). Prespeech vocalizations of a deaf infant: A comparison with normal metaphonological development. *Journal of Speech and Hearing Research, 28*, 47–63.

Oller, D. K., & Steffans, M. L. (1993). Syllables and segments in infant vocalizations and young child speech. In M. Yavas (Ed.), *First and second language phonology.* San Diego, CA: Singular Press.

Olmstead, D. (1971). *Out of the mouths of babes.* The Hague: Mouton.

Osberger, M., Robbins, A., Berry, S., Todd, S., Hesketh, L., & Sedey, A. (1991). Analysis of the spontaneous speech of samples of children with cochlear implants or tactile aids. *American Journal of Otology, 12*(Suppl.), 151–164.

Oyama, S. (1985). *The ontogeny of information.* Cambridge, UK: Cambridge University Press.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

Plotkin, H. C. (1997). *Evolution in mind: An introduction to evolutionary psychology.* Cambridge, MA: Harvard University Press.

Redican, W. K. (1975). Facial expressions in nonhuman primates. In L. A. Rosenblum (Ed.), *Primate behavior: Developments in field and laboratory research.* New York: Academic Press.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.

Steger, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature, 388*, 381–382.

Teixeira, E. R., & Davis, B. L. (1999, July). *Early sound patterns in the speech of two Brazilian Portugese Speakers.* Paper presented at the International Conference on Language Acquisition, San Sebastian, Spain.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Tyler, A. A., & Langsdale, T. E. (1996). Consonant-vowel interactions in early phonological development. *First Language, 16*, 159–191.

Vaivre-Douret, L., Le Normand, M.-T., & Wood, C. (1995). Developmental outcome in a tracheostomized child: A case study. *Early Child Development and Care, 112*, 13–26.

Vihman, M. M. (1978). Consonant harmony: Its scope and function in child language. In J. H. Greenberg (Ed.), *Universals of human language: Vol. 2. Phonology.* Palo Alto, CA: Stanford University Press.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications.* Timonium, MD: York Press.

Vihman, M. M. (1996). *Phonological development: The origins of language in the child.* Oxford, UK: Blackwell.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A reassessment of the continuity issue. *Language, 61*, 397–445.

Watson, J. B. (1925). *Behaviorism.* New York: Norton.

Zlatic, L., MacNeilage, P. F., Matyear, C. L., & Davis, B. L. (1997). Babbling of twins in a bilingual environment. *Applied Psycholinguistics, 18*, 455–471.