Functional organization of speech across the life span: A critique of generative phonology[1]

PETER F. MACNEILAGE AND BARBARA L. DAVIS

Abstract

Generative (Chomskyan) linguistics has a number of conceptual problems impeding its acceptance into modern cognitive science. A major issue centers on Uniformitarianism, the assumption that the supposed innate core of language – Universal Grammar – has remained the same since its instantaneous origin and remains basically the same (exhibits "continuity") throughout ontogeny. Other problems, noted by George Miller, are emphasis on structure rather than function, on competence rather than performance, as well as the tendency to regard simplifications as explanations. We illustrate these problems in the domain of speech acquisition, which, rather than exhibiting continuity, involves a progression from a syllable reduplication mode to an opposite syllable variegation mode. We present an alternative Neodarwinian conceptualization – the Frame/Content theory – in which the time domain is central for both phylogeny and ontogeny. According to this theory the reduplication-to-variegation progression in the ontogeny of speech (from syllabic Frames to segmental Content) is considered to recapitulate its phylogeny.

1. Introduction

There has been an unfortunate interdisciplinary boundary between much of modern generative linguistics and other branches of cognitive science. One main reason for this boundary is the core assumption of Chomskyan linguistics regarding the origin of linguistic phenomena. According to this perspective, both the syntactic and the phonological components of languages are considered to have evolved more or less instantaneously and not subsequently

1. This work was supported in part by NICHD RO1 HD27733-11.

The Linguistic Review 22 (2005), 161–181 0167–6318/05/022-0161

changed in any important ways. This claim is designated as Uniformitarianism (Newmeyer 2002). Not only is phylogenetic change denied, but in the ontogenetic domain in which change is undeniable, it is accorded a minor role, subsidiary to a basic "continuity" (Macken 1995: 674). Change is considered minor because the preordained form – Universal Grammar – provides the main structure for development, and any change is attributed to the process of setting language-specific values for preordained parameters. As Chomsky has described it, what happens in development is simply "triggering and shaping" of innate structure (Chomsky 1966).

This conception is incompatible with the most influential contemporary theory across a number of scientific disciplines regarding the nature of life forms – the Neodarwinian theory of evolution by natural selection. A central tenet of Neodarwinian theory is that every evolutionary development since the origin of the simplest life forms has been based on descent with modification. Instantaneous de novo evolution is thus proscribed. Linguistic capacities must, to some degree at least, have evolved by descent with modification from nonlinguistic capacities as linguistic output involves new uses of systems originally evolved for other purposes. And it is simply a matter of common sense that language must have initially been simpler than it is now, which means that a time-domain conception is essential to its phylogeny. In addition, if language did evolve gradually, it would not be surprising that the evolutionary sequence of events might be, at least to some degree, reflected in ontogeny, given that phylogeny involves a succession of ontogenies.

Regardless of one's perspective, the question of which of these two views of language is preferable is, ideally, empirical. In the absence of a fossil record for language, ontogeny is the most available site for empirical study. Unfortunately, as pointed out by Braine (1994) among others, the Chomskyan conception has no direct implications for the actual sequence of events in ontogeny at either the syntactic or the phonological levels. There have been three versions of universal grammar. None of them include a detailed conception of what the ontogenetic sequence of events should be. For example, in the Principles and Parameters version of Universal Grammar, (Chomsky 1981) a finite list of parameters has never been presented for syntax or phonology, let alone a time-domain-based conception of when they should be set in ontogeny – a conception of what exactly should develop and in what order.

In addition to the absence of an explicit basis for refutation which would be provided by a time-based conceptual structure, there appears to be an insuperable problem for the Chomskyan perspective in the area of phonological development. We have noted that the Chomskyan developmental structure is basically one of shaping of structures – setting of parameters – without introduction of new forms. However one of the main characteristics of speech acquisition is that infants prefer to follow a syllable with an identical one (termed redupli-

cation). In contrast, languages (in this context, mature speakers of languages) prefer to follow a syllable with a different one (variegation). One purpose of the present narrative is to illustrate why this problem (termed "infant-adult asymmetry" – Macken 1995: 673) *is* a problem for generative phonology, and to show how the problem might be soluble in the context of a Neodarwinan theory of speech evolution.

The problems that generative linguistic theory raises for cognitive science are not confined to the uniformitarian assumption. George Miller (1990), often deemed the father of cognitive science, identifies three other aspects of linguistics that have created communication problems between linguists and cognitive scientists. The first is the emphasis on *structure* rather than function. The second and related problem is the distinction made between competence and performance. The third, and most important in Miller's opinion, is that "Linguists tend to accept simplifications as explanations." (1990: 321): "For example, a grammarian who can replace language-specific rewriting rules with X-bar theory and lexicalization feels he has explained something: … For a psychologist on the other hand, explanation is something phrased in terms of cause and effect …" (1990: 321).

As to the first problem, the most systematic treatment of conceptual tension between structural (i.e., formal) and functional approaches is the monograph by Newmayer (1988). With respect to phonology, for example, Studdert-Kennedy (e.g., Studdert-Kennedy 1998) has repeatedly chided linguists for their characterization of distinctive features as descriptive predicates, unrelated to function. With respect to the competence-performance distinction for example Jackendoff (2002: 33) sees Chomsky's conception of competence as leaning towards what he calls a "hard idealization" which "denies the need to go beyond itself." (See also Jackendoff 1988.)

Of the three problems Miller identifies, the first two are longstanding problems of linguistics. As Braine has pointed out, in considering linguistics in general: "The goal of linguistic theory is to describe human languages. Simplicity and generality are gained by a description that … account(s) for as much *structure* (italics ours) as possible by means of principles and facts that are universal to all languages." (Braine 1994: 14) Linguistic *structure* has always been the prime concern of linguistics, and competence, or "knowledge of language" is a matter of language structure in Chomsky's view. While the competence-performance distinction was introduced in the middle of the twentieth century by Chomsky (1965), it was preceded by a similar distinction between "langue" (form) and "parole" (substance) made by Saussure (1915–1959) in which language *structure* was the domain of langue, and parole was actual speech.

Prior to the Chomskyan era, linguistics was primarily concerned with description, not explanation. The third problem – the problem of simplifications as explanations has primarily arisen in modern Chomskyan generative linguis-

tics as a result of the postulation of the instantaneously evolving innate generative grammar in the mind of every individual. Since then, it has become common to conclude that various regularities in linguistic structure, particularly those that are universal, are *explained* by innate Universal Grammar. As Braine puts it, "The assumption that such universals have an innate basis is a natural one." (Braine 1994: 14).

But however natural it might be in the context of generative linguistics, this assumption has perhaps been responsible for the preponderance of criticism of linguistics in the broader cognitive science community. The attitude of generative linguists toward this assumption and towards those who question it is perhaps best illustrated by Jackendoff in a paper entitled "Why are they saying these things about us." (Jackendoff l988) Why did he write this revealing paper? "Often I have found myself depressed by their apparent failure to recognize the value of the sort of work my friends and I do." (1988: 435)

Jackendoff asks: "What does phonology tell us about the mind, anyway?" He answers:

> If there is any message that I would hope nonlinguists would derive from our work, it is the hypothesis that one's cognitive and behavioral repertoire is significantly shaped by biologically transmitted *innate* (italics ours) knowledge, and that this knowledge can be expressed by a set of formal principles. In the larger picture, the importance of phonology and syntax comes from their status as the two domains in which this hypothesis has been most extensively investigated. (1988: 437)

In short, as Jackendoff sees it, the core contribution of linguistics is the hypothesis of innate form.

But Jackendorf finds that

> this too goes unappreciated. I have been asked by one of the major figures in cognitive science, a prominent philosopher of the Boston cognitive science community, Why do you people keep using this bogus appeal to innateness? You're getting yourselves off the hook too easily. (1988: 437)

The problem for Jackendoff, and perhaps for many other linguists, seems to be a lack of realization that describing something as being innate is not explaining it.

As already mentioned, one purpose of this narrative is to present a critique of the uniformitarian assumption underlying generative child phonology, focusing on the inconsistency of the assumption with the facts of infant/adult asymmetry – infants reduplicate syllables, adults variegate them. The second purpose is to provide a critique regarding generative phonologists' treatment of the crucial step when an infant begins to move from one side of the infant/adult asymmetry to the other; that is, when she begins to move systematically from syllable

reduplication to syllable variegation. The critique will identify instances subject to Miller's three criticisms of linguistic approaches.

In both of these contexts we will offer a Neodarwinian alternative to the linguistic approaches in the form of the Frame/Content theory of evolution of speech. This theory, we believe, avoids problems of linguistic approaches by providing a viable explanation of both the infant/adult asymmetry, as well as the first step the infant takes to begin to join adults.

2. Child phonology and the infant/adult asymmetry

As Macken points out in the chapter on "Phonological Acquisition" in the Handbook of Phonological Theory (Macken: 1995), the core assumption of generative child phonology is that the end state of speech acquisition is built-in in advance of its use. Consequently, a cornerstone of child phonology is what Macken calls the "Strong Identity Hypothesis" – "the capacity of children and adults is the same" (Macken 1995: 673). She outlines the situation in this way:

> In nature we find two contrasting development relations. In a relation manifesting essential continuity, the young are unskilled and simpler, yet they are fundamentally like the adults of the species in key respects; in a qualitatively different kind of developmental relation (which we might call "nonlinear") there is a radical difference between the beginning and end states and a major discontinuity in development.
>
> We are interested here in the nature of those changes: are they qualitative in the sense that the basic structures and capacities change, or quantitative in the sense that the information or knowledge of a specific domain changes. If the principles and objects of phonology are present at the outset of language learning, and thus instantiated at each stage, and in each interim grammar constructed by the learner, as in Chomsky's theory, then the developmental model is one of basic continuity. We would then look to nonqualitative factors to explain the developmental stages. If, on the other hand, some phonological principles or objects are not present at the outset, then there is no necessary relationship between a developmental stage of the child and the properties of phonological systems. The developmental model will then be one of discontinuity, and we explain the qualitative characteristics of each stage in terms of the maturation of new linguistic skills or changes in other cognitive capacities, as presented in Piaget's theories. (1995: 674)

The basic question then is whether the acquisition of the phonological component of language is characterized by continuity or discontinuity. It is well known that in babbling (7 to 12 months) and the 50 word stage of early speech,

(12–18 months) infants tend to repeat (reduplicate) the same syllable. In contrast, languages tend to variegate successive syllables – that is; make either the consonant or the vowel or both different in a second syllable than in a preceding one. (The term "reduplication" in the child phonology literature is used only to denote repetition of the same syllable at the phonological level, with no implications for morphology. This is different than the use of the term in language description where it is typically applied to participants in a morphophonemic alternation that are signaled by repeating a syllable.). We have gathered quantitative evidence for this tendency in consonant-vowel (CV) syllables with stop consonants and nasals, the most common consonants at these early stages of babbling and first word use. In a study of pairs of consecutive consonant-vowel syllables in 6 babbling infants (Davis and MacNeilage 1995), we found a median tendency to repeat the same syllable about 50 % of the time. The chance expectation for syllable repetition, computed on the basis of the relative frequencies of the particular consonants and vowels participating in the individual syllable pairs was 32 %. Chance expectation was determined in the following way: If a particular CV syllable type constitutes .3 of the instances of syllables in a corpus of pairs of adjacent syllables, then the probability that this syllable type will be reduplicated in a syllable pair is $.3 \times .3 = .09$. An expected proportion for reduplication of every syllable type in the corpus can be computed in this way. The sum of all the individual proportions is the overall proportion of reduplications to be expected. This is then compared with the actual proportion of reduplications – .5 in the above case. Thus reduplication occurred over 1½ times as often as expected by chance. In contrast, when we made a comparable analysis of pairs of consecutive CV syllables containing stops and nasals in dictionary counts of 10 diverse languages (MacNeilage et al. 2000a), we found that the same CV syllable was only repeated 2/3rds as often as would be expected by chance.

To our knowledge this problem of infant/adult asymmetry for phonological theory was first raised by Drachman (1976) in a criticism of the relatively short-lived school of Natural Phonology (Stampe 1973). As Anderson (1985) noted, the basic claim of natural phonology was that "what is natural in the systems of adult languages has its basis in whatever more general aspects of our phonetic capacities operate in other domains (such as child language …)" (Anderson 1985: 345). Drachman focused on the repetition of the same *consonant* in successive syllables in infants, a phenomenon called "Consonant Harmony", and pointed out that although this is pervasive in infant speech, it is virtually absent in languages.

The problem is still with us. In recent years there has been much optimism that a new approach to linguistic theory called Optimality Theory (Prince and Smolensky 1997) would help to solve long-standing problems of phonology. However, Pater (Pater 1997: 246) observes that "chief among" the "puzzles"

that "remain to be addressed … is the fact that although consonant harmony is extremely common in child language, it is unattested in this form in adult languages …"

As an attempt to deal with the presence of early consonant harmony in the context of Optimality Theory, in which sets of "Constraints" take center stage, Pater takes the position "that the constraint *Repeat* is constructed by the child in response to the pressures imposed by the developing production system. …" Pater's position involves the assertion that "at least some constraints of child phonology are inductively learned rather than innately given." (1997: 246–247). This attempt to deny innate status to consonant harmony is a peculiar one for two reasons. First it is difficult to see how the infant would *learn* harmony, when the ambient language models she is exposed to tend to avoid it. Second, as consonant harmony is dominant in infants from the very beginning of prespeech babbling, several months before the first words are produced, it seems to be giving the infant too much credit to say that she "constructs" it to solve problems of speech. If it was constructed, there should be a prior stage with less harmony before the new more mature phase occurs. Instead, infants *begin* with a reduplicative phase in which there are not only specific consonant harmonies that tend to be accompanied by patterns of vowel harmonies. Pater's conclusion is that "Clearly the introduction of child-specific constraints has implications for learnability theory that cannot be taken lightly. Not only would constraint reranking have to be shown to be computationally tractable but an account would also have to be given of constraint genesis and of constraint extinction." (1997: 247).

In short, even in the most contemporary approach to phonological theory in which a much more flexible approach to constraints on phonological form has been adopted, the infant/adult asymmetry remains a major stumbling block. We have developed an alternative Neodarwinian approach to phonological theory that includes a possible explanation of Pater's constraint genesis and constraint extinction. According to our construction of the problem space infant/adult asymmetry is exactly what would be expected. We will now outline the theory.

3. The frame/content theory

The subject of the Frame/Content (F/C) theory is the evolution and acquisition of the serial organization of speech, focusing on the syllable (MacNeilage 1998; MacNeilage and Davis 1990, 2000). It begins with the contention that modern adult speech has a Frame/Content (F/C) mode of organization shown by segmental serial ordering errors, in which consonants and vowels are misplaced in otherwise correct utterances. In these serial ordering errors, segments ("Content") obey syllable structure ("Frame") constraints whereby consonants

move into consonantal positions and vowels into vowel positions in syllables. For example, in reversal errors or spoonerisms, syllable initial consonants reverse with each other (*billfold → fillbold*), vowels reverse with vowels (*silly putty → sully pity*) and final consonants reverse with final consonants (*deep throat → deet thrope*) (Shattuck-Hufnagel 1979; Levelt l989). Most importantly vowels (syllable nuclei) and consonants (syllable margins) never adopt each other's positions in syllable structure in any segmental serial ordering error.

According to F/C theory, the frame constraint on speech errors is a premotor or cognitive constraint. This constraint evolved from a motor basis in mandibular oscillation (MacNeilage, l998). As the consonantal closing movement and the vocalic opening movement of the mandibular cycle have always been antagonistic to each other, there has never been a time when the premotor or cognitive underpinnings of the two movement phases could have gotten mixed up with each other. This assertion is a cornerstone of the embodiment perspective characterizing this theory. Mental structure is considered to have evolved in the context of constraints imposed by bodily action capabilities (Davis and MacNeilage 2000).

The mandibular cycle underlying the syllable, which we call the "Motor Frame", is considered to have evolved from the mandibular cyclicities of ingestion (chewing, sucking, licking) in paleomammals, circa 200 million years ago. In an intermediate stage, the cycle was exapted for visuofacial communicative cyclicities, (e.g., lipsmacks), forms common in other primates, before being paired with phonation (voicing) to form protosyllables. These hypotheses are elaborated in MacNeilage (1998). The recent finding of mirror neurons in the monkey homolog of Broca's area, which are activated by producing and observing ingestive cyclicities, and lipsmacks (Ferrari et al. 2003), in one instance, in the same neuron, provides strong neural evidence for this proposed phylogenetic sequence of events. Mirror neurons are neurons which discharge when an animal performs a specific action and also when it observes another animal performing the same action. Such neurons are possible precursors to a system of perception-production relationships such as that necessary for speech to be learned. It is particularly significant for F/C theory that the homolog of Broca's area and its surround is the main site of cortical control of ingestive processes in all mammals (MacNeilage 1998).

The initial protosyllables (e.g., "bababa"), presumably characterized simply by multicyclical reduplication because of their heritage in multicyclical ingestive actions and subsequently multicyclical smacks (e.g., "bababa"), could have been a component of the vocal grooming that Dunbar (1996) believes might have succeeded manual grooming when hominid group sizes increased to the point where hands-on grooming became impractical. Thus vocalizations could serve as an affiliative signal like tactile contact does in grooming but could be

given simultaneously to several animals, thus enhancing group solidarity. The ability to learn to make these protosyllabic forms in early language contexts may have been an offshoot of the evolution of a general-purpose mimetic capability, suggested by Donald (1994). Mirror neurons, which were originally found for manual actions, (Rizzolatti and Arbib 1998) before being found for ingestion and visuofacial communication would have been part of the neural substrate for the evolution of this capability.

What light does the acquisition of speech cast on F/C theory? The joint incidence of the consonant-vowel (CV) syllable as the only universal syllable in languages (Bell and Hooper l978) and the ubiquity of the CV form in the babbling and early speech of infants is one of a number of reasons to ask whether the study of speech ontogeny might throw light on phylogeny.

As suggested by our initial findings (Davis and MacNeilage l990), we and others have shown in a number of studies in various language communities (summarized in Davis and MacNeilage 2002) that, co-existing with the mandibular cycle, there are severe inertial constraints on the internal structure of CV sequences in babbling and early words. For example, the tongue tends to stay in a similar position for the consonant and the vowel. Three types of consonant-vowel co-occurrence patterns have been found: (a) Coronal or tongue front consonants co-occur with front vowels, (e.g., [daedae]); (b) dorsal or back consonants co-occur with back vowels (e.g., [gogo]), although both dorsals and back vowels tend to be rare at these stages; (c) and labial or lip consonants co-occur with central vowels, which presumably involve a resting position of the tongue (e.g., [baba] or other central vowels perhaps differentiated in height only by mandible height). We refer to this overall pattern as "Frame Dominance" (Davis and MacNeilage l995) because all these forms involve frames in the absence of any other active articulator placement during the utterance. The early word stage differs from the babbling stage in that a focus on attempts at particular lexical items often results in CV patterns other than the three predicted ones (e.g., [no] in the infant studied by Davis and MacNeilage 1990). But nevertheless, at a statistical level the three basic patterns continue to be typically observable (Davis and MacNeilage 2002). These patterns are also typically present in languages. This has been shown in both our work (MacNeilage et al. 2000) and that of Rousset (2003) who analyzed a total of 24 diverse languages. Kinney and MacNeilage (2002) have also reported these patterns in 9 Creole languages.

The same biomechanical constraints limiting tongue movement from consonant to following vowel (CV) also apparently limit it from vowel to following consonant (VC) because the same 3 co-occurrence patterns are found in this sequence as well (MacNeilage and Davis 2002). However in languages where, we have seen, there is constraint against repeating the same syllable, we found that the biomechanical constraints on vowel to consonant movement (across

syllable boundaries) were absent in languages (MacNeilage et al. 2000). They were also absent in the 9 Creole languages (Kinney and MacNeilage 2002).

The absence of these patterns in some languages makes it possible to decide whether infants are simply copying their patterns from the ambient language, or whether they are inherent in early production system operation. The latter seems to be the case. For example the labial-central preference has been firmly established in English-speaking infants, but no such preference is shown in a dictionary count of English (MacNeilage et al. 2000). Lee, Davis and MacNeilage (2005) have also shown that all three preference patterns are shown in Korean infants even though none of the preferences exists in the Korean language. And in unpublished work we have found that in the first word stage, infants unsuccessfully attempting to produce a non-preferred sequence tend to err in the direction of producing the preferred form.

What is the meaning of these findings for the relationship between speech phylogeny and ontogeny? First, the finding that the three CV co-occurrence constraints observed in infants tended to be present in languages as well suggests that the same biomechanical factors that limit infant CV production probably limited early hominid CV production as well. Why else would they be characteristic of languages? One could perhaps argue that these co-occurrence tendencies were only introduced later in language when on-line pressures associated with successful production of complex sequences led to sound changes that favored a least effort principle. But if these patterns only developed later, then the earlier simpler patterns produced by the same biomechanical system are unexplained. Our thesis is that because biomechanical constraints on the oral mechanism predated speech, they must have influenced the earliest speech.

But while infants also showed VC co-occurrence patterns corresponding to the CV patterns, languages did not. Why is the same set of biomechanical constraints present in languages in one context (CV) and not the other (VC)? This suggests that just as infants eventually surmount their biomechanical constraints on vowel-consonant sequences, thus achieving independent control of successive syllables in learning modern languages, earlier hominids eventually surmounted these initial constraints in phylogeny. They presumably did this in the course of evolving independent control of individual syllables, in response to selection pressures to increase the size of the linguistic message set.

At present F/C theory does not systematically extend to the phylogeny and ontogeny of more complex syllable shapes. This is perhaps not a great limitation at our present stage of understanding, partly because complex syllable shapes are relatively rare in languages. Maddieson (1999) has concluded from a survey of a large number of languages that 80 percent of languages are restricted entirely or virtually to the CV form. Within generative phonology, an innate mental basis for syllable structure is considered to exist in the form of a "Sonority Principle". It is considered to be an innate basis for *all types of*

syllable, regardless of their complexity (e.g., Blevins 1995). According to this principle, syllable nuclei (vowels) are the loudest sounds and loudness tends to taper off as a function of distance from the vowel in consonants on either side of it. But in fact the surface forms of sonority can be attributed primarily to the biophysical fact that the more open the mouth is, the more efficiently sound is projected. There is no evidence that there is a mental counterpart to this phonetic property which is innately available.

Our assertion based on present data is that syllable structures more complex than CVs will eventually be explicable in terms of a trade-off between the tendency towards favoring articulatory ease and the continuing demand for perceptual distinctiveness. Here again ontogeny may repeat phylogeny, progressing from exhibiting more to exhibiting less biomechanical constraint. Consistent with this possibility, Jakielski (1998) has shown that like CV structures, infants' early consonant clusters are subject to Frame Dominance. In a study of 4 infants between the ages of 7 and 35 months she found that 5/6ths of the clusters were homorganic (shared roughly the same place of articulation), and could therefore be simply a result of minor modulations of a single mandibular elevation movement without marked changes in place of articulation. From this result we would predict that homorganicity in clusters in languages will be inversely proportional to cluster inventory size and complexity. To our knowledge, generative phonology makes no predictions about the relation between cluster structure and cluster frequency in languages.

To conclude, the Neodarwinian Frame/Content theory offers an explanation for the ontogenetic progression from reduplication to variegation, a progression which cannot be dealt with by the generativist continuity theory of speech acquisition. Part of the reason the Frame/Content theory can handle the issue of explanation is that, unlike uniformitarianism, the descent with modification tenet allows the possibility that while some important things can remain basically the same (in this case the frame) other things can change, (in this case content) giving the impression of an overall discontinuity. But as Stephen J. Gould observed: "external discontinuity may be inherent in underlying continuity provided that a system displays enough complexity." (Gould 1977: 409). To the argument that such changes are not inconsistent with Chomsky's "triggering"metaphor, we would reply that the connotation of shaping here is one of setting of parameters *already provided*, and a reversal of the basic pattern of segmental serial organization from reduplication to variegation goes beyond such a connotation.

4. The transition from reduplication to variegation: Generative approaches

There is some agreement that the first systematic change an infant makes in going from reduplicating sequences of syllables to variegation, thus showing increased intersegmental control, is to begin to produce CVC or CVCV words in which the first consonant in the word is a labial, and it is followed after the vowel by a coronal. We will call this the *LC effect*. This effect seems to only develop in the first word stage. We have not found it in babbling (Davis and MacNeilage 1995). An example of such a word is [bado] for *bottle*. In a review of 7 reports involving 5 different language communities we found that 21 out of 22 infants preferred the LC sequence (MacNeilage et al. 2000b). In addition, in our own work, we found the effect in 9 out of 10 infants in the 50 word stage in data available from other reports (MacNeilage et al. 1999).

The strength of this tendency is indicated by reports that infants sometimes produce this form even when the target word has a CL structure. Macken (1995) noted a strong tendency of one Spanish-speaking infant to prefer the labial-coronal sequence even when the word target he was attempting had a coronal-labial sequence ([pwaeta] for [sopa]). In addition, Lee (2003) has found that Korean infants show a preference for LC forms even though the Korean language does not.

Let us now consider the question of how this initial step in the progression from favoring reduplication to favoring variegation is treated in generative child phonology. Clartje Levelt (1994), a generative phonologist, has accumulated the most data relevant to this question. Levelt concludes that reduplicative forms are a result of Distinctive Features being specified for the entire phonological word. (Distinctive features are sub-attributes of consonants and vowels.) For example, in the instance in which an infant produces [puf] for Dutch *poes* 'cat' the joint feature of labial consonant – rounded vowel is considered to apply to the entire word. Then when the infant subsequently produces the LC pattern [pus] – the correct phonetic rendering of Dutch *poes* – Levelt suggests that the "left edge" of the word "becomes available for separate place of articulation specification". Thus, in this example the labial specification becomes restricted to the initial (left edge) consonant and the following vowel. Levelt considers that the final coronal is then produced as a result of "Underspecification". This assumption requires some explication.

Underspecification has become an important concept in phonology. To take a simple example, it is considered unnecessary to specify the segment /s/ in the input for English words which have 3-consonant onset clusters (e.g., *screw*, *splice*) as all such words begin with an [s]. Consequently [s] can be inserted at a late stage of the derivation of such words as a default specification. (A derivation is a set of conceptual steps leading from an underlying representation –

the input – to a surface representation – the output.) Coronals are the most favored consonants for underspecification. They are considered to be the most *unmarked* consonants (Paradis and Prunet l991). The concept of *Markedness* pertains to the complexity and relative frequency of linguistic units and structures. "Unmarked" designates minimal complexity and maximal use (Prince and Smolensky 1997).

The motivation for the use of the underspecification concept in phonology is that it increases the formal simplicity of derivations – it is a simplification. But from the point of view of those interested in the causes of speech – immediate or proximal causes rather than ultimate causes in this case – underspecification is an extremely counterintuitive concept. In concrete terms, it involves the existence of empty slots in the representation of words in the mental dictionary, and raises the question of how an infant could acquire mental representations of words which are partially empty at an underlying level. However, the empirical question, to which we do not have an answer, is how the underspecification assumption could be successfully tested.

Levelt regards her characterization of the LC effect as preferable to earlier generative proposals in that it results from "a co-occurrence restriction on consonants to the effect that the first consonant in the word must have a more front place of articulation than the second one." (Levelt 1994: 75) This view, called "Fronting", was originally formulated by Ingram (l974) who concluded that the phenomenon was an instance of markedness.

For Levelt, this characterization of the LC effect in terms of a left edge effect on distinctive features and an underspecified final coronal consonant characterized as unmarked is considered to be an explanation of the form, sufficient to, in her words "account for" its production. She also considers distinctive features and the phenomenon of underspecification to be innate mental entities. However distinctive features and underspecification are simply entities of structural description which have not been related to the performance level – the level of the actual movements the infant makes to produce this form. The concept of markedness, which we also encountered earlier in Pater's exposition suffers from two major problems. First, the notion of complexity has not been adequately defined. Defining markedness in terms of relative frequency involves circular reasoning – more frequent is designated as more unmarked, and then markedness is considered to explain relative frequency (e.g., Menn 1983). (We have already considered the problem that reduplication is more frequent, and therefore more unmarked in infants but variegation is more frequent, and therefore more unmarked in adults – an impossible outcome for a single markedness dimension.) Thus, Levelt's characterization of the LC effect has all three of the problems alluded to by Miller.

Macken (1995) has also given an account of the LC effect. She considers that both the production of consonant harmonic and non-harmonic consonant

sequences such as LC involve consonant-consonant *templates*, which she calls "harmony" and "melody" templates respectively. Each of Macken's templates apparently involves a vowel slot between the two consonants which can be filled with *any* vowel. This involves the assumptions that vowels and consonants are more independent of each other at this early stage of speech than they are. It does not take into account the fact that there are co-occurrence constraints between consonants and vowels. Macken (1995) did not believe that infants showed co-occurrence constraints between consonants and vowels, even though they had already been reported in a number of studies (Jakobson 1968, Stoel-Gammon 1983, Davis and MacNeilage 1990).

Again, as in the case of Levelt, the concepts provided – templates – are structural entities, not related to performance. Though they were basically redescriptions of sound patterns, they were considered to explain the behaviors in question in terms of innate entities, in this case part of what Macken called "autonomous phonology". To postulate two types of template, harmony for reduplication and melody for variegation is a completely ad hoc solution to the problem posed by this momentous step. It does not address the question of why one template tends to be used before the other, or the question of how this sequence of events represents continuity.

Velleman (1996) has analyzed metatheses (sound reversals), such as the one produced by Macken's Spanish-speaking infant, from the perspective of Optimality Theory. Using this approach, she was able to characterize not only typical patterns such as the LC pattern, but also other non-typical patterns of individual infants reported in the literature. Velleman concluded that "[a]bsolute claims about metathesis such as fronting [Ingram l974], and strict positional restrictions [Levelt l994, Macken l995] are too inflexible for the full range of patterns which occur. Varying patterns of metathesis … can all be *accounted for* (italics ours) using a constraint-based approach, within which specific manner and place features are aligned with edges of words." (Velleman 1996: 183). Once again, "accounted for" is intended to mean "explained" but no consideration is given to *why* these patterns occur.

All three of these generative approaches have the problems noted by Miller. They are concerned with structure rather than function, competence rather than performance and involve descriptive simplification rather than explanation. None of them address the issue of how infants actually *do* the LC pattern, or *why* that particular pattern is favored.

5. The frame/content theory and the LC effect

How is the LC effect to be explained in the context of the Frame/Content theory? First, one additional fact needs to be added. The LC effect is also present

in languages. In our survey of 10 languages (MacNeilage et al. 2000b), we observed it in 9/10. It was also observed in Rousset's (1993) corpus of 14 languages, and in Kinney and MacNeilage's (2002) corpus of 9 Creole languages. It is an extremely robust effect. The mean ratio of LC forms to CL forms in the 10 language sample was a remarkable 2.25:1 which compares favorably with a median ratio of 2.5:1 in our 10-infant database.

First of all, as described earlier, we explain the underlying tendency to alternate between consonants and vowels in general (not satisfactorily addressed by the generative theorists) in terms of the evolution of the motor frame. In our view, the LC effect involves a specific modification of organizational patterns present in babbling (MacNeilage and Davis 2000). There are two main frame types in babbling and early speech. One is what we call "Pure Frames" involving labials and central vowels (e.g., [bababa], or [mamama]. We call them pure because we believe they are produced entirely by mandibular action without active tongue movement. The other type involved coronals and front vowels (e.g., [daedaedae] or [naenaenae]. We call this pattern "Fronted Frames" because they require a fronting movement of the tongue superimposed on what we believe to be the same mandibular oscillation cycle as for pure frames. (As we have said, the third type, dorsal-back or "Backed Frames" is often absent or rare in babbling and early speech.)

We consider pure frames to be easier to produce than fronted frames because the latter require an additional tongue placement movement superimposed on a frame common to labial-central and coronal-front forms. There are other reasons to believe that pure frames are easier. There are 3 studies in the literature that show that infants who have tracheostomies (a hole in the trachea into which a plastic breathing tube is inserted) up to the age of two and have therefore not been able to vocalize normally, strongly favor labial consonants after the tracheostomy has been sealed (studies are cited in MacNeilage and Davis 2002). In addition, it has been shown that the frequency of labial forms in general increases in first words relative to babbling. For example Boysson-Bardies et al. (1992) showed it in 4 languages. We interpret this change as a regression to easier output forms in the presence of the new additional functional load associated with the new need to link output forms with lexical concepts.

Based on this evidence, we propose that LC forms begin with the pure frame because it is easier. Then instead of a tongue fronting movement occurring before the utterance begins as in the usual reduplicated fronted frames, the fronting movement is made to coincide with the beginning of the *second* frame cycle.

Why should the easier form occur first? It is well known in motor system neurophysiology that initiating movement requires a special operation involving separate circuitry in addition to what is required if movements are to be continued (MacNeilage and Davis 2001). In certain clinical syndromes (e.g.,

Parkinson's Disease) there are specific problems with movement initiation. We believe that the tendency to favor LC over CL forms is a self-organizational consequence of the interaction of these factors (relative ease, the new lexico-motor functional load component, and the demands of movement initiation) in the context in which an infant is attempting to produce a variegated form in the language.

This hypothesis includes predictions for what vowels are likely to occur in L-vowel-C sequences. As central vowels co-occur with labial consonants and front vowels with coronal consonants, most vowels would be expected to be in the central or front dimensions but not the back. Information relative to the expected identity of vowels in L-C contexts is available from an intensive case study of one infant (MacNeilage and Davis l999). The observed-to-expected ratios of the three vowel types in coronal environments were front, .96; central 1.28; back, .31 (1.00 is a chance value). These observations, from data gathered before the hypothesis was formulated, strongly support the self-organizational hypothesis. They are also counter to Levelt's scenario, as rounded vowels in Dutch are back vowels, but back vowels tend not to accompany labials in this infant's LC forms.

What is the phylogenetic significance of the LC form? Like the CV co-occurrences, we have pointed out that the LC effect is characteristically present in languages. We found it in 9 of 10 languages (MacNeilage et al. 2000) The exception was Japanese. We interpret its presence in languages to mean that it was fundamental to the creation of variegation in the history of languages and was probably present at early stages of phylogeny as a response to the need to increase the size of the set of distinguishable lexical forms. We hypothesize that the same variables that lead to production of this form as a self-organizational response in infants were present in earlier hominids with one exception. Instead of having variegated forms as the linguistic model that they must match as modern infants do, earlier hominids invented this form with the aid of self-organizational forces. That is, in the presence of the relative ease of pure frames, the demand to interface the lexicon with the movement control system and special demands of movement initiation hominids tended to produce the LC form. Therefore it tended to be available as a sound pattern to which a concept could be linked.

6. Conclusions

We have noted some problems with the generative approach to language, particularly to speech acquisition which impede the acceptance of the discipline into mainstream cognitive science. We have offered an alternative Neodarwinian Frame/Content approach to speech, specifically speech acquisition

which we feel lacks these problems. We conclude with some observations regarding the implications of (F/C) theory for the relation between the body and the mind, and for the concept of innateness.

A possible criticism of the F/C theory is that with its emphasis on motor determinants of speech it is denying that there is a significant mental level of speech representation. Not at all. We assert that the nature of mental organization of speech derived from the necessity that it be compatible with the properties of the motor apparatus with which it is associated. Note that this view is in contrast with Chomsky's assumption in a monograph with the significantly formalistic title *The Architecture of Language* (2001) that a mental phonological level evolved first and that it either just happened to be compatible with sensorimotor capabilities or was followed by a mutual accommodation of the two components. What would such an initial unconstrained mental apparatus have been like considering that phonology is meaningless outside of the context of communication, and why did it take the form that it did? We believe that to the degree that the present-day mental apparatus derived from the necessity to be compatible with motor constraints, we must reconstruct the derivational process to the contemporary phonological endpoint. We do not deny the possibility that there could be purely mental aspects of speech, but simply assert, following Miller, that those who believe there are such aspects need to give a causal explanation of them.

To illustrate the possibility that mental aspects of speech can be derived from aspects of bodily processes without assuming innateness, we briefly present the following example. In a number of neurological syndromes, patients produce speech automatisms in the form of reduplicated CV syllables (e.g., "babababa", MacNeilage and Davis 2001). These automatisms are apparently produced primarily in medial premotor cortex – the Supplementary Motor Area (SMA) and the Pre-SMA. We argue that the existence of these infantile forms in adults is not simply a matter of regression but an indication of the existence of a *cognitive-motor* frame production capability which helps interface the movement control system with the mental lexicon. We suggest that this partially mental frame normally lies behind speech production but is in this case, conveniently for phonological science, produced in isolation. But although this is probably a universal mental aspect of speech it is not necessarily innate. It could be the epigenetic result of normal experience. As open-close alternations at the motor level are a constant and integral part of speech from the beginning of babbling onwards, a developing speaker may come to form a mental representation of this constant component as she develops a complementary capacity to go beyond reduplication and variegate output forms at the segmental level. The result is that the frames become programmable at a level higher than direct action, the premotor level. This analysis suggests that even universal mental aspects of speech need not necessarily be

innate, but may instead "fall out" of more basic aspects of speech experience.

It should be emphasized that we are not saying that structural descriptions with their attendant simplifications, and competence models are necessarily problematical. The field of linguistics has distinguished itself by providing characterizations of its subject matter at the structural and competence levels that are perhaps more valuable than those in any other area of cognitive science. For example, we could presumably profit from an analogous taxonomic analysis of the other great action system – the manual system. But from a Neodarwinian standpoint it is functions and performance that are subject to selection and therefore, in the long run, what must be understood if we are to explain language in terms of ultimate causes. To return to Miller's relatively orthodox conception of explanation, we would assert that generative phonology has not yet provided a causal explanation of *any* phenomenon in its discipline in terms of *why* the phenomenon is the way that it is rather than some other way. From this standpoint, assumptions of innateness in generative phonology, based at present only on simplifying descriptions, need to be accompanied by hypotheses as to how the phenomena became innate in orthodox evolutionary terms, and by a theory that is constrained by assumptions about exactly how the supposed innateness should play itself out in speech acquisition. From this perspective, the F/C theory is not simply a notational variant of generative phonological approaches plus an ad hoc phylogenetic component, derived from the very different and therefore possibily irrelevant domain of ontogeny. It is a fundamentally different way of constructing phonology that recognizes from the outset the necessarily integral relation between phylogeny and ontogeny.

The recent monograph by Jackendoff (2002) goes a long way towards a solution of the problems of linguistic theory discussed here. He attempts to integrate the findings of mainline descriptive linguistic theory with an impressive coverage of allied disciplines more concerned with performance-related aspects of language and with language evolution. He emphasizes the importance of an incremental (basically descent with modification) scenario for language evolution in preference to uniformitarianism, and even endorses our notion of protosyllables as an early component of language evolution (2002: 242–244).

We see this work as an important step towards the unification of cognitive science with linguistics. However, Jackendoff still takes a somewhat uncritical approach to the innateness concept. We have found no need for speech-specific innateness in the Frame/Content theory. The Frame has a deep evolutionary heritage, and the ability to learn to speak may have been part of a more general evolving mimetic capacity deriving from mirror neuron systems. The content component may have become increasingly complex in response to pressures to preserve perceptual distinctiveness while *inventing* a larger and larger lexical message set which tended to take the production system further and fur-

ther away from its most basic biomechanical modes. This development can be conceived in terms of units of imitation which Dawkins (1976) has called "Memes" which can replicate themselves in culture in something like the way genes replicate in nature.

University of Texas at Austin

References

Anderson, Stephen R. (1985). *Phonology in the Twentieth Century*. Chicago: The University of Chicago Press.

Bell, Alan, and Joan B. Hooper (1978). *Syllables and Segments*. Amsterdam: North-Holland.

Blevins, Juliet (1995) The syllable in phonological theory. In *A Handbook of Phonological Theory*, John Goldsmith (ed.), 206–244. Oxford: Blackwell.

Boysson-Bardies, Benedicte, Marilyn Vihman, Liselotte Roug-Hellichius, Catherine Durand, I. Landberg, and F. Arao (1992) Material evidence of infant selection from the target language. In *Phonological Development*, Charles Ferguson, Lise Menn, and Carol Stoel-Gammon (eds.), 369–392. Parkton, Md.:York Press.

Braine, Martin D. S. (1994). Is nativism sufficient? *Journal of Child Language* 21: 9–31.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press.

— (1966). *Cartesian Linguistics*. New York: Harper and Row.

— (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

— (2001) *The Architecture of Language*. Oxford: Oxford University Press.

Davis, Barbara L. and Peter F. MacNeilage (1990). The acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research* 33: 16–27.

— (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research* 38: 1199–1211.

— (2000). An embodiment perspective on the acquisition of speech perception. *Phonetica* 57: 229–241.

— (2002). The internal structure of the syllable. In *The Evolution of Language out of Prelanguage*, Talmy Givon and Bertram F. Malle (eds.), 135–154. Amsterdam: John Benjamins Publishing Company.

Davis, Barbara L., Peter F. MacNeilage, and Christine L. Matyear (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches. *Phonetica* 59: 75–107.

Dawkins, Richard (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Donald, Merlin (1994). *Origin of the Modern Mind*. Cambridge, Mass: Harvard University Press.

Dunbar, Robin I. M. (1996). *Grooming, Gossip and the Origin of Language*. Cambridge, Mass: Harvard University Press.

Drachman, Gaberell (1976). Child language and language change: A conjecture and some refutations. In *Recent Developments in Historical Phonology*, Jacek Fisiak (ed.), 123–144. The Hague: Mouton.

Ferrari, Pier Francesco, Vitore Gallese, Giacomo Rizzolatti, and Leonardo Fogassi (2003). Mirror neurons responding to the observation of ingestive and communicative mouth movements in the monkey ventral premotor cortex. *European Journal of Neuroscience* 17: 1703–1714.

Gould, Stephen J. (1977). *Ontogeny and Phylogeny*. Cambridge, Mass.: Bellknap.

Ingram, David (1974). Phonological rules in young children. *Journal of Child Language* 1: 49–64.

Jackendoff, Ray (1988). Why are they saying these things about us? *Linguistic Inquiry* 6 (3): 435–442.

— (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jakielski, Kathy Jo (1998) Motor organization in the acquisition of consonant clusters. Unpublished Ph.D. Dissertation, University of Texas at Austin.

Jakobson, Roman (1968). *Child Language, Aphasia, and Phonological Universals* The Hague: Mouton.

Kinney, Ashlynn, and Peter F. MacNeilage (2002). Fundamental properties of phonological organization: Consonant-vowel relations in Creoles. Paper presented at the Fourth International Conference on the Evolution of Language, Harvard University, March, 2002.

Lee, Sue-Ann (2003) The phonetic basis of early speech acquisition in Korean. Unpublished Ph.D dissertation, University of Texas at Austin.

Lee, Sue-Ann, Barbara L. Davis, and Peter F. MacNeilage (2005) Frame dominance and the serial organization of babbling and first words in Korean infants (Ms).

Levelt, Clartje (1994). The *Acquisition of Place*. Leiden: Holland Institute of Generative Linguistics.

Levelt, Willem J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MIT Press.

Macken, Marlys (1995). Phonological acquisition. In *A Handbook of Phonological Theory*, John Goldsmith (ed.), 671–696. Oxford: Blackwell.

MacNeilage, Peter F. (1998). The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences* 21: 499–546.

MacNeilage, Peter F. and Barbara L. Davis (1990). Acquisition of speech production: frames, then content. In *Attention and Performance XIII: Motor Representation and Control*, Marc Jeannerod (ed.), 452–468. Hillsdale N.J.: Erlbaum.

— (1999). Consonant (vowel) consonant sequences in early words. *Proceedings of the 14th International Congress of Phonetic Sciences*, vol. 3, 2489–2492. San Francisco: University of California, Berkeley.

— (2000). On the origin of the internal structure of word forms. *Science* 288: 527–531.

— (2001). Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology* 11: 696–700.

MacNeilage, Peter F., Barbara L. Davis, Christine L. Matyear and Ashlynn Kinney (1999). Origin of speech output complexity in infants and in languages. *Psychological Science* 10: 459–60.

MacNeilage, Peter F., Barbara L. Davis, Ashlynn Kinney, and Christine L. Matyear (2000). The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Development* 71: 153–163.

Maddieson, Ian (1999) In search of universals. *Proceedings of the 14th International Congress of Phonetic Sciences*, vol. 3, 2521–2528. San Francisco: University of California, Berkeley.

Menn, Lise (1983). Development of articulatory, phonetic, and phonological capabilities. In *Language Production Vol. 2. Development, Writing, and Other Processes*, Brian Butterworth (ed.), 4–50. New York: Academic Press.

Miller, George A. (l990). Linguists, psychologists, and the cognitive sciences. *Language* 66: 317–322.

Newmayer, Frederick (2002). Uniformitarian assumptions and language evolution research. In *The Transition to Language*, Alison Wray (ed.), 359–375. Oxford: Oxford University Press.

Paradis, Carole and Jean-Francois Prunet (eds.) (1991). *The Special Status of Coronals.* San Diego: Academic Press.

Pater, Joe (1997). Minimal violation and phonological development. *Language Acquisition* 6: 201–253.

Prince, Alan, and Paul Smolensky (1997). Optimality: From neural networks to universal grammar. *Science* 275: 1604–1610.

Rizzolatti, Giacomo and Michael Arbib (1998). Language within our grasp. *Trends in Neuroscience* 21: 188–194.

Rousset, Isabelle (2003). From lexical to syllabic organization: Favored and disfavored co-occurrences. *Proceedings of the 15th International Congress of Phonetic Sciences*, vol. 1, 715–718. Barcelona: Autonomous University of Barcelona.

Saussure, Ferdinand, de (1915). *Cours de Linguistique General*. (English translation: C. Bally and A. Sechehaye (eds.)) (1959) *Course in General Linguistics*. New York: Philosophical Library.

Shattuck-Hufnagel, Stefanie (1979). Speech errors as evidence for a serial ordering mechanism in speech production. In *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, William E. Cooper and E.C.T. Walker (eds.), 295–342. Hillsdale, NJ: Lawrence Erlbaum Associates.

Stampe, David (1973). The acquisition of phonetic representations. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*. 443–454.

Stoel-Gammon, Carol (1983). Constraints on consonant-vowel sequences in early words. *Journal of Child Language* 10: 455–457.

Studdert-Kennedy, Michael G. (1998). The particulate origins of language generativity: From syllable to gesture. *In Approaches to the Evolution of Language*, James R. Hurford, Michael Studdert-Kennedy and Chris Knight (eds.), 202–221. Cambridge: Cambridge University Press.

Velleman, Shelley (1996). Metathesis highlights feature-by-position constraints. In *Proceedings of the University of British Columbia International Conference on Phonological Acquisition*, Barbara Bernhardt, John Gilbert, and David Ingram (eds.), 173–184. Somerville: Cascadilla Press.