that children may have less-detailed representations of the sound structures of lexical items.

However, there is also evidence from studies of vocabulary growth that even young children do not avoid learning homonyms and do not seem to be confused by them (E. V. Clark 1993; Landau and Shipley 1995; also see the arguments in Gerken, Murphy, and Aslin 1995). Homonyms seem to lie at the most extreme end of confusability of sound patterns of words (i.e., they are identical). Nevertheless, there may be a difference between having two items with the same sound pattern, and two items that are only minimally different in their sound patterns. Ultimately, this issue will only be resolved through studies that systematically chart the growth and organization of the lexicon.

We will return to the issue of lexical organization in the next two chapters. The aim here was only to point out some possible arenas in which we might expect to see constraints pertaining to sound properties interact with those pertaining to meaning.

Chapter 7

Relating Perception to Production

Acquiring a spoken language requires becoming not only a fluent perceiver but also a fluent producer of speech. At the same time that infants are listening to and learning about speech sounds produced by others around them, they are beginning to produce speech sounds of their own. Their productions of speech must eventually conform to the sounds that they hear in the native-language input. Yet, to this point, we have been considering developments in speech perception without regard to what is happening in speech production during the same time frame. Whether it truly makes sense to study the development of speech perception independently from the development of speech production is debatable. On the one hand, there are theories that argue that speech production is intimately involved in the process of speech perception, either because decoding the signal requires reference to the movements used to produce it, as in motor theory (Liberman et al. 1967), or because articulatory gestures are directly perceived in the speech signal, as in ecological theories (Best 1995; Fowler 1986; Studdert-Kennedy 1991b).

If one adheres to either of these views, the core of perception and production is the same—motor plans in one case, gestures in the other—in which case it makes the most sense to study the development of the two systems together. On the other hand, there is the fact that different peripheral physiological devices are involved in these activities—the articulatory system, for production, and the auditory system, for perception. Thus, although the speech signal provides a common ground as the output of production and the input to perception, differences in the organizations of the articulatory and auditory systems raise the possibility that the systems may develop more or less independently of each other. Of course, the degree to which the two systems are inextricably linked or independent in
their development is a question that can only be answered by empirical research.

This chapter reviews some of the important findings regarding the development of speech production. As with the earlier discussion of speech perception capacities, I focus on changes in speech production that relate to the acquisition of a native language. Much as my review of speech perception capacities began by focusing on their relation to phonetic distinctions and then moving on to a consideration of how these develop to support word recognition in fluent speech, I follow a similar course with respect to the development of speech production. Hence, after a brief description of Roman Jakobson's influential views on phonological development, I discuss babbling and how it develops in accordance with speech input, and then move toward the production of words and larger utterances. Along the way, wherever possible, parallels between significant changes in production and perception are noted. Some discussion of the development of the lexicon as it relates to speech production is also included. The chapter concludes with an examination of the means by which production and perception are related in the course of acquiring a native language.

Jakobson's Views on Phonological Development

Jakobson (1941) was responsible for the first, and best-known, theory of phonological development. He proposed a structuralist account of the way children acquire the sound structure of their native language. He distinguished between two periods of phonological development: babbling (which he believed to be prelinguistic) and the true acquisition of language sound structures. Jakobson did not actually collect any data regarding children's speech productions. Rather, he based his arguments on observations reported by Greigore (1933), who had claimed that babbling behavior is "prelinguistic" and not "the first genuine stage of language." Jakobson also endorsed the view that during the prelinguistic stages, the child would produce a large variety of different sounds that were not found together in a single language or even a group of languages.

A child during his babbling period can accumulate articulations which are never found within a single language or even a group of languages—consonants of any place of articulation, palatalized and rounded consonants, labials, affricates, clicks, diphthongs, etc. According to the findings of Greigore (1937), the child at the height of his babbling period "is capable of producing all conceivable sounds." (Jakobson 1941, 21)

This claim bears certain parallels to the more recent claims in the speech perception literature (see chapter 3) that infants begin speech perception with the capacity to discriminate contrasts in any of the world's languages. However, unlike Jakobson, most contemporary speech perception accounts see a continuity between these early perceptual capacities and subsequent language-specific perception. By comparison, Jakobson believed that following the babbling period in which infants were producing this large variety of sounds, there was a discontinuity that occurred between babbling and speech.

As all observers acknowledge with great surprise, the child then loses nearly all of his ability to produce sounds in passing over from the pre-language stage to the first acquisition of words, i.e., to the first genuine stage of language. (Jakobson 1941, 21–22)

Jakobson sometimes referred to an intervening silent period between these two phases of development (see also Velten 1943). However, at other times, he seemed to suggest that the silent period was not a necessary step in the developmental sequence.

As Meumann has already stated (1903, p. 23), a short period may sometimes intervene between the stage of spontaneous babbling and that of true language development in which children are completely mute. For the most part, however, one stage merges unobtrusively into the other so that the acquisition of vocabulary and the disappearance of the pre-language inventory occur concurrently. (Jakobson 1941, 29)

In any case, the true beginnings of language acquisition were hypothesized to commence only at the end of the babbling period. Thus, babbling came to be seen as at best a preparatory stage for language production, but something short of true linguistic behavior (e.g., Fry 1966).

Jakobson believed that linguistic universals, principles that hold across all languages, were the major determinants of how infants learned language. Although there might be individual differences in the rate of acquisition, each child was hypothesized to go through an orderly series of stages in mastering the full range of phonemic contrasts that appeared in the native language. The sequence of these stages was determined by what he called "laws of irreversible solidarity," which were basically an inherent universal hierarchy of structural laws. A phonological system for a given language was stratified, or layered, and this ordering of layers was taken to be universal across languages, and therefore invariable. Ferguson and Garnica (1975) offer a concrete example of this sort of relationship:
"No language has nasal vowels unless it also has one or more nasal consonants." This universal property of language structures would then translate into a prediction about acquisition of nasal consonants and nasal vowels, namely, children should acquire nasal consonants before they acquire nasal vowels. More generally, because of this stratified arrangement, the discovery of a certain type of phonemic contrast within the phonological system might depend on first discovering some other type of phonemic contrast, depending on how the two types are ordered. Another way to put this is that, if the phonological system has a contrast of type Y (e.g., nasal vowels in the example above), this would necessarily imply that the system also has a contrast of type X (e.g., nasal consonants). That is, the occurrence of Y is dependent on the occurrence of X in the system.

One interesting aspect of Jakobson's views was that he expected these kinds of universal principles to apply to all aspects of linguistic behavior. Thus, when language abilities degenerate because of aphasia or some other language disorder, the layers that are added last in development should be the first to disappear, whereas the ones that were developed first should be the last to go. Moreover, whenever reacquisition is necessary, such as after a stroke that results in language loss, then the course of reacquisition must follow the original developmental sequence. Because the universal order of phonemic development that Jakobson proposed was derived from his inspection of different languages, he also assumed that phonemic contrasts that were relatively rare across languages would be ones that were most likely to be acquired late in development (and consequently, the first lost by aphasics).

The sequence of stages was hypothesized to proceed from a simple undifferentiated state toward an end state that was stratified and highly differentiated. He proposed that children begin with the optimal vowel (i.e., a wide vowel such as [a]) and the optimal consonant (i.e., a labial stop, such as [p]). The pairing of these two elements in succession in a single syllable, serves to define the basic syllable structure as CV. From this point, differentiation occurs with the production of other syllables with this same basic form (e.g., from "papa" to "mama" or "tata"). The component segments of these new syllables differ with respect to certain distinctive features from those in the first CV syllable. These distinctive features pertained primarily to the manner in which sounds are articulated (e.g., whether they are voiced or voiceless, whether they are produced with nasal resonance or not, whether there is complete or incomplete closure of the vocal tract, whether there is a secondary source of noise present or not) and with the resonance characteristics of the oral cavity (e.g., whether the closure or narrowing of the vocal tract during production occurs in the front, middle, or back of the oral cavity). With respect to the sequence in which phonemic oppositions are acquired, after the initial oral-labial stop, the child was usually expected to acquire a nasal consonant next—an observation that subsequently led Jakobson (1971) to suggest this as the reason why parental names in various languages are likely to begin with these kinds of consonants.

Jakobson's theory had a great impact on views about how children acquire the sound patterns of their native language. His views served as a reference point for many other investigators who gathered data about speech production and phonological development.

**Babbling as the Beginning of Speech Production**

A factor that has worked to divorce research on the development of perception from that on production is that most data relating to these two domains were collected with infants at different ages. Many speech perception studies focused on infants 6 months of age or younger, whereas most of the data collected on speech production was with infants well beyond their first year. This picture has changed during the past decade or so, in part because many more studies of speech production have more carefully documented the development of babbling and its relation to early word productions. Certainly, improvements in the technology available for recording and analyzing babbling have helped to encourage such studies.

Aside from the sheer technical problems, another factor tended to hold back research on babbling behavior during the first year. This had to do with the legacy, derived from Jakobson's writings, that babbling has little to do with phonological development as it relates to word production. The view that there was a discontinuity between babbling and true language acquisition was reinforced by reports about the behavior of deaf infants. Lenneberg (1967) noted that the onset of babbling in deaf infants appeared to occur at about the same age as in their normally hearing counterparts. Consequently, babbling was deemed to arise independently of any experience in hearing language input. By comparison, there was no indication that deaf infants actually began to produce words at the age when normally hearing infants did. Thus, this pattern of results was seen as another indication that babbling is prelinguistic.
The explosion of studies on phonological development from the late 1970s onward has yielded much information about the relationship between early babbling and productions of the first words. The evidence from these investigations challenges Jakobson's claims regarding the discontinuity between babbling and word production and casts doubt on the existence of an intervening silent period (for extensive reviews see Locke 1983; Vihman 1996; Vihman and Elbert 1987). Contrary to the notion of a silent period, a survey of many previous investigations reveals that many of the children studied continued to babble even after they began producing their first words (Blount 1969; Elbers 1982; Ferguson and Farwell 1975; Labov and Labov 1978; Leopold 1947; Olmsted 1971).

Moreover, the phonetic characteristics of early words often are continuous with patterns prevalent in the child's babbling. Elbers and Ton (1985) have commented on the fact that "words and babble are formed from the same pool of productive sound patterns." Similarly, other investigators (Stoel-Gammon and Cooper 1984; Vihman et al. 1985; Vihman and Miller 1988) have reported that the repertoire of sounds and sound combinations used by any given child in late babbling and in early word production are closely related. Thus, not only is there no evidence of a silent period, but there is little support for the notion of a discontinuity in the sounds that appear in babbling and those that occur in the production of the first words. Therefore, rather than accept Jakobson's claims that babbling is prelinguistic, it is more reasonable to treat babbling behavior as a genuine stage of language development.

Similarly, the assumption that deaf infants babble normally has been further scrutinized and found wanting. Gilbert (1982) criticized the nature of the evidence on which Lenneberg had based his claims. Furthermore, several recent investigations with deaf infants who were more carefully screened with regard to the nature and extent of their hearing losses have yielded findings that are inconsistent with Lenneberg's claims. For example, in their study with hearing-impaired and normal infants, Stoel-Gammon and Otomo (1986) found that hearing-impaired infants produced utterances with fewer different consonant types per session. Moreover, in contrast to normally hearing infants, the hearing-impaired infants showed a clear decrease in the production of different consonant types over sessions.

A longitudinal investigation with hearing-impaired infants by Oller and Eilers (1988) produced evidence even more damaging to Lenneberg's claims. Oller and Eilers obtained extensive auditory screening measures for their subjects. They noted that by comparison to normally hearing infants who begin canonical babbling (i.e., producing strings of alternating consonants and vowels) at between 6 and 10 months of age, the hearing-impaired infants did not enter this stage until much later, 11 to 25 months of age. Furthermore, several months after the onset of this stage of babbling, only a third of the hearing-impaired infants produced vocalizations that were judged to be in the range of ones that normally hearing infants produced. Consequently, these findings suggest that babbling does not develop normally in the absence of auditory input.

Similarly, the available evidence suggests that Jakobson's claim about the diversity of sounds present in babbling is also in error. Locke (1983) points out that data collected with 12-month-olds in studies by Irwin (1947) and by Pierce and Hanna (1974) indicate that the 12 most frequent consonants that occur in babbling account for 95 percent of all the consonants heard in the infants' productions. By comparison, the 12 least frequent consonants in babbling, which were largely fricatives, affricates, and liquids, appeared in only about 5 percent of the infants' utterances. In addition, Cruttenden (1970) reported that two children whom he studied showed a number of important omissions with respect to the sounds included in their babbling repertoire. He concluded that contrary to Jakobson's claims, the subjects that he followed did not babble all the sounds of English. In fact, as data from more recent investigations based on acoustic analyses of babbling show, children's productions of particular speech-sound categories only roughly approximate those in any adult language (Boysson-Bardies 1993; Boysson-Bardies et al. 1989).

Thus, although infants are certainly gaining more control over their vocalizations during the babbling period, there is no reason to treat this phase of development as separate from what happens when infants actually begin to attempt to produce their first words. In fact, as the data from speech perception suggest, infants at the ages when babbling is prevalent (from 4 to 12 months and beyond), are very attentive to the sounds of language. Although the infants clearly are not speakers of any language, they are certainly no less involved in the process of acquiring a language than they are at a later age when producing their first words.

There is also evidence of a different kind that supports the notion that babbling is a natural and, likely, a necessary part of language acquisition. In view of the earlier discussion above, it is ironic that some of these data were obtained with deaf infants. Work by Petitto and her colleagues (Petitto 1993; Petitto and Marentette 1991) on the acquisition of sign
language suggests that learners of such languages also go through a period of babbling. Thus, in addition to the usual range of non-linguistic gestures that infants produce, infants who are exposed to sign language begin to produce a range of gestures that are signlike. Moreover, these productions are not limited to infants who have been deprived of spoken language input. Hearing infants who have been exposed to sign language because their parents are deaf begin to produce such "manual babbling" more or less simultaneously with vocal babbling. Thus, as with learners of spoken languages, learners of sign languages appear to go through a period in which they experiment in producing the kinds of elements that potentially could be combined to form words in their native language. Hence, these parallels in how production develops for units in sign languages and in spoken languages provide an additional reason for considering babbling behavior to be a true part of the development of speech production.

Developmental Changes in Babbling

Having made the case for babbling as a part of language development, let us consider what is known about the development of babbling. Vocalizations prior to 6 months of age are often confined to isolated vowel-like sounds that are sometimes accompanied by nasalization or glottal consonants and velar/uvular fricatives (Roug, Landberg, and Lundberg 1989). An important development occurs at around 6 months of age, when the infant begins to engage in so-called reduplicated (or canonical) babbling (Oller 1980; Roug et al. 1989; Stark 1980; Vihman 1993a). Reduplicated babbling is said to occur when the child begins to produce sequences of syllables, usually consisting of a stop consonant combined with an open, central vowel (e.g., "babababa"). There is little variation in either intonation or constituent consonant and vowel segments. Similar tendencies have been noted for such languages as English (Oller 1980; Stark 1980), Dutch (Koopmans van Beinum and van der Stelt 1986) and Swedish (Roug et al. 1989). Oller (1980) has argued that this period is the first point at which the child produces syllables that could potentially serve as building blocks of words.

Following this period in which reduplication of syllables is prevalent in infants' vocalizations, babbling becomes more complex in ways that approximate utterances from the native language. This period has been referred to as variegated babbling. Vocalizations are now characterized by an alternation of consonantal segments rather than a reduplication of a consonant (Roug et al 1989). The production of variegated babbling overlaps with reduplicated babbling. However, the proportion of variegated babbling seems to increase considerably around 12 to 14 months when the child begins to produce a variety of consonants overlaid on a sentence-like intonation pattern. It is during this time that word production is beginning and a marked increase occurs in the number of word patterns that the infant attempts to imitate (Vihman and Miller 1988).

A number of general trends in the way that babbling behavior develops among infants from different language backgrounds have been ascribed to anatomical and physiological changes that occur throughout the first year. The infant's vocal apparatus undergoes many changes during this period of time. For example, the newborns' vocal tracts are considerably shorter, their tongue shapes are different, and their larynges are in a much higher position than those of adults (Lieberman 1984). Kent and Miolo (1995) suggest that until about 3 months of age, infants' vocal-tract configurations are more similar to those of nonhuman primates than to those of adult humans, and only at 4 months of age do infants' vocal tracts begin to resemble those of human adults. However, development of the vocal tract continues well after this point. Crelin (1987) has observed that the descent of the larynx to its eventual adult locus only begins during the third year. Because of the immature configuration of their vocal tracts, it is difficult for young infants to produce sounds corresponding to the vowels [i] and [u] (Buhr 1980; Lieberman 1977; Lieberman, Crelin, and Klatt 1972). A general tendency that has been noted in babbling behavior is an evolution from a rather centralized vowel space to one that is considerably more spread out at the end of the first year (Buhr 1980; Kent and Murray 1982). In addition, there is evidence that early on in development there is a predominance of front vowels relative to back vowels (Lieberman 1984).

Other regularities in babbling that have been observed appear to be related to infants' ability to gain control and coordination over jaw movements. For example, tendencies for particular consonants and vowels to occur together with greater frequency in babbling than in the adult target language led MacNeilage and Davis (1990) to suggest that these kinds of co-occurrence patterns might have a common frame that was produced by the oscillation of the jaw. This, the frequent association of labial consonants and central vowels was explained as the result of lip closure produced by the oscillation of the jaw combined with the tongue
remains in its resting position. Similarly, they found evidence for the
association of front consonants with front vowels, which they attributed
to the absence of active changes of tongue position during oscillation.
Indeed, MacNeilage and Davis (1991) have even offered an explanation
for variated babbling that is based on variation in the amplitude of jaw
movements during successive oscillations.

Davis and MacNeilage (1994) also reported evidence for the interaction
of segmental and suprasegmental properties in the productions of an
English-learning infant whom they studied between 7 and 12 months of
age. In particular, one unexpected finding was that when stress increased
from one syllable to the next, the vowel tended to shift from central to
front (and in the opposite direction—from front to central—when stress
decreased). This pattern is counter to the usual relationship found for
vowel quality and stress. Davis and MacNeilage suggest that one possible
reason for the association that their infant displayed is that the produc-
tion of front vowels (which involve moving the tongue forward) may be
easier to combine with the increased amplitude of jaw opening (which is
itself related to the production of stress). Davis and MacNeilage sug-
gested that this kind of association between stress and vowel type in early
productions may not be a universal tendency because many languages
make little or no use of stress. For this reason, it would be instructive to
know just how widespread this kind of relation is in the babbling behavior
of infants. Information about the behavior of infants exposed to lan-
guages that differ in their predominant word-stress patterns would help
resolve the question of the association that Davis and MacNeilage ob-
erved is universal or not.

In any event, the studies reviewed in this section indicate that, not sur-
prisingly, certain characteristics of infants’ babbling patterns stem from
the nature of their articulatory apparatus and the degree of control that
they have over its movements. We now turn to a consideration of the
effects that exposure to sound patterns of a specific language have on
infants’ vocalizations.

Effects of Linguistic Experience on Babbling

Given that reduplicated babbling is observed cross-linguistically, one can
ask about the impact of the native language on babbling. Recall that
Brown (1958) had suggested that babbling drifts in the direction of the
target language. Subsequently, Weir (1966) reported that analyses of her
data pointed to considerable differences in the babbling behavior of
American, Russian, and Chinese infants from 6 months onward. How-
ever, other researchers found little empirical support for babbling drift
toward the sound patterns of the native language. For example, Atkinson,
MacWhinney, and Stoel (1968) reported that adult listeners could not
reliably judge the language background of infants based on their bab-
bling. Similarly, Olney and Scholnick (1974) found no evidence that their
subjects could reliably distinguish samples from English and Chinese
infants ranging from 6 to 18 months. In contrast, Boysson-Bardies,
Sagart, and Durand (1984) reported that French adults distinguished the
babbling of a French infant from an Arab infant but not from a Chinese
infant.

However, using adult judgments of native and nonnative babbling
patterns is not the only way to evaluate claims that native-language input
affects babbling patterns. Collecting cross-linguistic data on infant vocal-
izations and subjecting these to acoustic analyses can provide a means of
assessing the effects of linguistic experience. In a survey of existing reports
of babbling from a variety of different languages, Locke (1983) found
little evidence of differences in phonetic repertoires related to language
background. He also remarked that there was considerable variation in
the phonetic repertoires of infants from the same language background.
Hence, there were large individual differences in the pattern of phonetic
acquisition among infants exposed to the same language input. On the
basis of these findings, Locke argued that there was little support for the
view that babbling drifts toward the native language at this phase of
development.

Much of the data that Locke reviewed relied exclusively on investiga-
tors’ phonetic transcriptions of infants’ babbling patterns. Better audio
recording techniques and acoustic analysis tools have helped to improve
the quality and reliability of judgments regarding the phonetic patterns
found in infants’ babbling. Investigations using these new techniques
combined with larger cross-linguistic sample sizes have provided empiri-
cal support for the claim that babbling drifts toward native-language
sound patterns. For example, Boysson-Bardies et al. (1989) noted that
vowel production in 10-month-olds from four different language back-
grounds (Parisian French, Algerian Arabic, British English, and Hong
Kong Cantonese) tends to parallel differences found in adult productions
of vowels in these languages. In particular, the formant structure of
vowellike patterns produced by infants from different language backgrounds differ from each other in ways that resemble productions of these vowels by adult speakers of these languages. Moreover, similarities have also been noted between infants’ babbling and adults’ productions of consonants. An extensive longitudinal study (Boysson-Bardies, Vihman, Roug-Hellichius, Durand, Landberg, and Arao 1992) with infants from French, Swedish, English, and Japanese backgrounds plotted the drift toward native-language consonantal categories. The infants were seen from 9 months of age until they achieved vocabularies of 25 words or more. The distributions of consonants in infants’ babbling and in their early word productions were compared to the sample of adult words that served as targets in the language that the infants were acquiring. Even at 10 months, there were differences in the distribution of manner and place categories in the vocalizations of infants from the different language backgrounds. For example, analyses of the distribution of stop consonants in adult production indicated that these were most prevalent in Swedish followed by English, then Japanese and, lastly, French. The production of stops by the infants showed exactly the same pattern. Thus, there are indications that the phonetic segments that appear in babbling from 10 months of age onward are influenced by the target language.

Native-language influences on babbling occur not only for phonetic properties but also for other aspects of sound pattern organization. For instance, there is evidence for early language-specific prosodic influences (Levitt 1993). In a study of five French- and five English-learning infants between 5 and 13 months, Whalen, Levitt, and Wang (1991) found international differences in babbling that were consistent with those in the adult target languages. Also, rhythmic properties relating to the timing of syllables in the native language may emerge even earlier than any influences on segmental properties of babbling (Levitt, Utman, and Aydelott 1992).

Just as studies in the perceptual realm indicate that 9-month-olds are attentive to the way that segments are typically ordered and combined in their native language (e.g., Friederici and Wessels 1993; Jusczyk, Friederici, et al. 1993), investigations of syllable structure in babbling also reveal effects of language-specific influences toward the end of the first year. Levitt et al. (1992) found that between 11 and 13 months, their French-learning infant was less likely to produce closed syllables than was their English-learning infant. This difference is in line with the distribution of open and closed syllables in the two target languages.

Vihman (1992) also found evidence for differences in syllable structures among infants from four different language backgrounds (Swedish, Japanese, English, and French). In particular, although all the children she studied had certain phonetic tendencies in common (producing syllables such as [da], [ha], and [hal]), there were also cross-linguistic differences that were in line with tendencies found in the input language. For instance, Japanese infants were much more likely to produce velars in association with back vowels (e.g., [ko], [go]). In addition, Swedish infants produced syllables in which [t] and [d] were associated with a full range of vowels, rather than just front vowels, as infants from the other language backgrounds tended to do.

Therefore, in addition to the constraints that are imposed on infants’ speech sound repertoires by the nature and maturation of their articulatory apparatus, the distribution and variety of sounds that they babble are affected to some degree by the characteristics of their target language.

Relating Babbling to Word Production

Until fairly recently, much of the research on phonological development focused on the production of words by language learners. Such studies generally examined the numbers and kinds of words that children have in their productive vocabularies (e.g., Ferguson and Farwell 1975; Menn 1978; Moskowitz 1970; Smith 1973). In noting the growth and development of the child’s productive vocabulary, many of these approaches have emphasized the progress that language learners make toward producing the full range of phonemic oppositions in their native language (Ferguson and Garnica 1975; Ingram 1974a, 1978; Macken 1980a; Menn 1978; Stampe 1969). The descriptions provided in many of these early studies are stated in terms of the acquisition of particular phonemic contrasts and phonological rules (Garnica 1973; Ingram et al. 1980; Menn 1980; Moskowitz 1970; Smith 1973). Hence, one can come away with the impression that the child’s focus at this stage is on learning new phonemic contrasts, rather than new words (although see Ferguson 1986; Ferguson and Farwell 1975; Waterson 1981 for views that whole words may constitute the initial representations).

From the point of view of a linguist interested in when a child might give evidence of having a particular contrast, these kinds of descriptions were certainly appropriate. However, it is important not to assume that the linguist’s description of what oppositions are appearing at a given age
is equivalent to what actually drives the child's behavior during acquisition. The child's goal at a given moment may not be so much to acquire a certain phonemic contrast as it is to successfully produce particular words (recall the parallel point that was made about speech perception in the discussion at the end of chapter 4). Indeed, changing the focus from the production of phonemic segments to the production of words may even help account for some of the puzzles in phonological development (such as the variability in the production of particular segments on different occasions or in different contexts) that Ferguson and Garnica (1975) noted in their review. In what follows, rather than addressing particular theories of phonological development, I confine the discussion to some important phenomena that have been noted in children's productions of words in their native language.

One important issue concerns the relation between the sounds that appear in babbling and those that occur in the first words that learners produce. For example, Viikman and her colleagues (Viikman et al. 1985) found that the phonetic characteristics of English-learning infants' early words are highly similar to those of their contemporaneous babbling. Similarly, in their investigation of the babbling monologues of a Dutch child during the period of the production of first words, Elbers and Ton (1985) reported mutual influences between babbling and word production. Babbling seemed to give rise to phonological preferences in the selection of word targets from the adult language, whereas word production also led to changes in the kinds of phonetic segments that appeared most frequently in babbling patterns. In a cross-linguistic investigation that involved infants from English-, French-, Swedish-, and Japanese-learning environments, Boysson-Bardies and Viikman (1991) also found some evidence for continuity between the characteristics of babbling and early word production. Language-specific patterns with respect to the production of consonants in babbling (i.e., asymmetries in terms of the frequencies with which certain manner or place classes occurred across the languages), also tended to show up in the early productions of words by these same infants. At the same time, Boysson-Bardies and Viikman noted that there were some interesting differences that occurred in the productions of some types of consonantal classes. For example, across all the languages, the percentage of stops was higher in word production than it was in babbling, whereas for fricatives the reverse trend occurred. Following Kent's (1992) hypothesis regarding avoidance of or substitutions for phonetic segments, they attributed these changes to production con-

strains. Specifically, they cited constraints that arise from the control required to produce fricative segments intentionally, which may have led to the substitution of stops for fricatives.

Findings such as these do point to continuities in babbling and in early word production. At the same time, they give an indication that things change when the infant moves from simply producing strings of sounds, or as Ferguson and Macken (1983) put it, "playing with sounds," to intentionally trying to produce words. Although the same articulators are involved in each instance, the task demands associated with trying to use sounds to convey a particular meaning appear to cause the child to fall back on articulatory routines that may be simpler to control. Perhaps these sorts of task demands are also a contributing factor to the use of vowel or consonant harmony in children's early attempts at producing certain words in the adult language (e.g., [gāk] when attempting to produce "duck").

Another frequently cited phenomenon in children's early productions has to do with the variability that arises in the production of the same word by the same child on different occasions. Perhaps the best-known case comes from Ferguson and Farwell (1975), who reported that the word "pen" received a number of phonetically different pronunciations by the same child in the course of a single observation period. This type of variability in pronunciation has been widely reported (Leonard et al. 1982; Menn 1976; Priestly 1976; Schwartz and Folger 1977; Waterson 1978). In addition, Locke (1983) has noted that many of the simplifications that appear in children's early words have been observed cross-linguistically (e.g., stopping, fronting, initial stop voicing, final obstructive devoicing, gliding, consonant harmony, cluster reduction, etc.). This suggests that variations in the input that the child receives are not likely to be a major factor in accounting for variability in production. For example, although it is conceivable that variations in input account for some small proportion of variability across children, this would not explain why the productions of a single child (like the one that Ferguson and Farwell observed) should be so variable. It seems far more plausible to attribute the variability to factors having to do with the infant's ability to control fine movements of the articulators. Indeed, studies have shown that infants are variable with respect to timing of motor movements (Kent and Murray 1982; Macken and Barton 1980). The pressures of having to produce an utterance whose meaning is correctly conveyed to the listener
could further reduce the capacity of the infant to control the timing of the relevant articulatory movements.

The suggestion that variability in production is not so much an inability to produce the required articulatory gestures as difficulty in producing them while trying to convey meaning is further supported by the prevalence of several other phenomena that have been noted. For example, there is the phenomenon that Gerken (1994a) has referred to as *chain shifts*. In such a shift, a child substitutes a different sound for one in an adult word (e.g., refers to “duck” as “guck”). This might lead one to conclude that the child has difficulty in producing [d] in word-initial positions. However, the same child might at the same time produce [d] in word-initial position as a substitute sound in a different word, such as saying “duck” to refer to “truck” (Macken 1980b; Menn 1978; Smith 1973; Stemberger 1992). Another situation that demonstrates that the child’s problem does not have to do with a general inability to make the required articulatory gestures is the case that Leopold (1939) first described with respect to his daughter Hildegard’s early word productions. Her correct pronunciation of the word “pretty” as “priti” was well in advance of her ability to produce consonant clusters in other words, which she routinely reduced to single segments. In other instances, a child’s favored, but inaccurate, pronunciation of some adult word might remain in the child’s vocabulary long after he or she is producing other words with the same target sounds correctly.

These findings suggest that when a child consistently produces an item in the same way, it may be because he or she has articulatory (motor) routines that are, more or less, rote associated with that particular word. By comparison, other words in the child’s vocabulary may not have specific, stored articulatory routines associated with them. Instead, the child may have to generate and sequence the correct articulatory routines online from a stored representation of the word’s sound pattern. If this view is correct, then variability in production of the same word stems from difficulty in handling the processing demands of generating the appropriate articulatory routines from a stored representation of the sound pattern while trying simultaneously to convey a specific message.

In addition to having particular articulatory routines that are associated with specific lexical items, learners may also have some general kinds of routines that they apply to many different kinds of items. For instance, the child may follow some general routine that allows only certain types of sounds to be produced in particular word positions, such as producing only voiced consonants in word-initial positions or producing fricatives only in word-final positions (Ingram 1974b; Kiparsky and Menn 1977; Menn 1980). Moreover, these general routines that the child employs can be quite idiosyncratic. Thus, Ferguson (1979) commented on the fact that a child may have a favorite sound that he or she uses repeatedly in the production of different lexical items. Vihman (1986) has suggested that the child may accidentally stumble across the correct production of a particular kind of sound, such as a liquid, in the course of articulatory exploration and then end up producing this sound in word productions more frequently, relative to other children at the same age. In fact, in her study of 10 English-learning children who were observed at weekly intervals between 9 and 16 months of age, Vihman (1986) noted a number of individual differences among her subjects in the diversity of segments used, the range of adult consonants that were targeted, the extent to which children continued to babble after they were producing 15 words, and the consistency and integrity of word shapes. However, when these same children were later seen at 36 months of age, many of these differences had receded. Vihman pointed to this fact as an indication of the unifying influences that native-language input has on the course of phonological development—a point that was further reinforced in the cross-linguistic data that she and Boysson-Bardies subsequently collected (Boysson-Bardies and Vihman 1991).

A number of investigators have pointed out that at least two distinct kinds of learning styles appear in the acquisition of phonology. Ferguson (1979) discussed one type of learner as a “cautious system builder” who constructs a tight phonological system. He characterized the other type of learner as a bolder child who shows a loose and variable phonological organization. Many others have found evidence that concurs with Ferguson’s observation (Bretherton, McNew, Snyder, and Bates 1983; Klein 1978; Menn 1978, 1983; Peters 1977). A recent paper by Peters and Menn (1993), which contrasts the ways two different children, Seth and Daniel, acquired grammatical morphemes in English, provides a further elaboration of the differences in these styles. Specifically, they note:

At least two strategies have been identified in the language acquisition literature: (1) The formulaic strategy (also called gestalt, expressive, or pronominal) where the focus is on multisyllabic chunks of speech and (2) the word-oriented strategy (also called analytic, referential, or nominal) where the focus is on shorter stretches (Peters 1977, 1983). From a prosodic point of view, it looks as though formulaic children, like Seth, pay initial attention to ‘horizontal’ information such as the
number of syllables, stress, intonation patterns (with only secondary attention to particular consonants and vowels); word-oriented children, like Daniel, pay more attention to the vertical segmental information contained in single (usually, stressed) syllables focusing on the details of consonants and vowels. (p. 745)

As Peters and Menn go on to suggest, the adoption of one of these two styles does not prevent children from learning the kind of information that is the focus of the alternative style. Eventually, children who follow either one of these two strategies end up learning the full range of facts concerning the organization of sound patterns in their native language. However, these children appear to take very different routes to acquire the same phonological ends.

Note that the categorization of children as being either formulaic or word oriented should not blind us to the fact that these are broad characterizations of learning styles and that there are many individual differences among those who might be assigned to either group. As Gerken (1994a) has remarked, the reasons for the appearance of these individual differences in acquisition strategies are not clear. Differences in the input to which the child is exposed may have something to do with these differences, but so far there is little empirical evidence to support this hypothesis (Leonard, Newhoff, and Mesalum 1980; Viilman 1993b). Another possibility is that these differences in style simply reflect temperamental differences among learners. Thus, it has often been suggested that some children may avoid attempting to produce words with sounds that they have not fully mastered in their phonetic repertoire (Ferguson and Farwell 1975; Macken 1978; Schwartz and Leonard 1982). These children may be conservative about making mistakes and thus end up as the kind of cautious system builders that Ferguson described.

Another possibility that has been mentioned is that differences in the perception and production systems of different children could influence the kinds of sounds that are most salient to them in the input and therefore most likely to be attempted by a given child (Studdert-Kennedy 1986). Along these lines, Viilman (1993b) has put forth the idea of an articulatory filter that is defined as “a phonetic template (unique to each child) which renders similar patterns in adult speech unusually salient or memorable.” Her suggestion is that language learners’ choices of which words to imitate are influenced by their knowledge of their own vocal motor schemes, so that patterns in the input that match those schemes are the ones most likely to be attended to and imitated. That is, sounds that learners hear frequently and that they know they can produce are more likely to capture their attention.

Whatever the reason for the individual differences in learning styles, there are still some generalizations that we can make about children’s first word productions relative to targets in the adult language. For the most part, when differences occur between targets and actual productions, the latter are apt to be simplifications of the adult words in terms of the numbers and kinds of phonetic segments that are produced (Ingram 1974b, 1978; Macken 1979; Menn 1978; Smith 1973). Gerken (1994a) describes several of the more commonly observed types of changes that children make. Substitutions occur whenever the child replaces a sound in an adult word with another one. Usually, the substituted sound is one with similar phonetic properties to the one replaced. In deletion, the child simply leaves out some of the phonetic material from the adult word. For example, the child may pronounce “banana” as [naen]. Metathesis occurs when the child reorders some of the phonetic material in the adult word such as in pronouncing “spaghetti” as [pazgeti]. Another simplification, cluster reduction, is one that we have already referred to. The child will often reduce consonant clusters by deleting one or more of the component consonants. Finally, assimilation is a process by which the child alters the pronunciation of one part of the word with a feature or phoneme from another part of the word. We considered an instance of this sort in describing the behavior of a child who pronounces the word “duck” as [gak].

Prosodic Characteristics of Children’s Early Productions

In addition to the phonetic composition, target words in the adult language have particular prosodic characteristics that also must be reproduced properly in order to ensure their correct recognition. Most investigations of phonological development, until relatively recently, have focused more on the phonetics than the prosody of children’s word productions. Still, some tendencies were noted in the prosody of children’s early productions. For example, Macken (1979) reported that when her Spanish-learning subject noticed that two words had the same rhythmic structure, she began to produce them both with an identical initial consonant. Elbers (1985) commented that whenever her subject had difficulty recalling an adult word, he substituted one with the same number of syllables and a similar stress pattern as the target word.
One prosodic characteristic of children's early word productions that has been frequently noted over the years is their tendency to omit unstressed syllables from adult target words (Bladell and Jensen 1970; Ingram 1974b; Smith 1973). This tendency is usually attributed to unstressed syllables being less perceptible than stressed ones in speech directed to the language learner (Brown and Fraser 1964; Chiat 1979; Echols and Newport 1992; Gleitman and Wanner 1982). However, as we noted then, the data from studies of speech perception abilities of infants suggest that they are capable, from a young age, of detecting phonetic differences between unstressed syllables of speech sounds (Jusczyk et al. 1978; Jusczyk and Thompson 1978; Williams 1977a).

Recent investigations by Gerken (1991, 1994b, 1996; Gerken et al. 1990; Gerken and McIntosh 1993) have suggested another explanation for the child's omission of unstressed syllables in early productions of both words and phrases. In particular, Gerken has argued that these omissions are attributable to constraints on production rather than perception. For example, she (Gerken 1994b) and others (Wijnen, Krikhaar, and den Os 1994) have pointed to the fact that children are more likely to omit unstressed syllables from some word positions as opposed to others as evidence against a perceptually based explanation for these omissions. In fact, these kinds of omissions are more likely to occur in unstressed word-initial syllables than in unstressed word-final syllables (Echols and Newport 1992; Gerken 1994b; Ingram 1974b; Vihman 1980). Thus, children are much more likely to omit the unstressed first syllable of "giraffe," than they are to leave out the unstressed second syllable of "monkey."

Gerken (1991) has argued that this pattern of unstressed syllable omission reflects the use of a particular type of production strategy on the part of early learners. In particular, she hypothesizes that the utterances that children produce at this early stage of word production are governed by a metrical production template that follows a trochaic pattern (i.e., an initial stressed syllable, followed by an unstressed one). Moreover, Gerken has shown that this kind of template can not only account for unstressed syllables in individual words, but it also accounts for which syllables the children are likely to leave out when they produce their first word combinations (Gerken 1994b; in press). For example, the use of a trochaic template for production helps to explain why children are prone to leave function words out of utterances, despite the fact that they apparently perceive them (Gerken and McIntosh 1993). However, another recent investigation by Boyle and Gerken (in press) showed that multiple factors are responsible for the omission of function words. In particular, the presence of unfamiliar lexical items in an utterance may itself be a sufficient reason for children to omit function words from sentences. Thus, across a series of three experiments in which 2-year-olds were asked to repeat sentences, they more frequently omitted function words in sentences that contained either novel nouns or novel verbs than they did for sentences with familiar nouns and verbs. This familiarity effect did not interact with whether the critical function word fit into a trochaic template. Thus, lexical familiarity and metrical properties appear to independently influence children's productions of function words.

Still, it is interesting that, just as the beginning stages of word segmentation in speech perception are characterized by the use of a strong/weak (i.e., trochaic) template, a similar tendency to impose a trochaic template also shows up in the child's early productions of meaningful words. We will return to this point shortly.

How Are Perceptual Representations Related to the Child's Productions of Words?

Another kind of behavior that appears in children's early productions has been termed the "fit" phenomenon (Berko and Brown 1960; Dodd 1975; Gerken 1994a; Smith 1973). A child may systematically mispronounce a particular word in a certain way (e.g., "fish" is pronounced as [fi]). One possible interpretation of this is that the child has somehow misperceived the adult word. Consequently, it is the perceptual representation on which the pronunciation of the word is based that is the source of the error. However, this explanation of the child's behavior is undercut by the fact that the child who makes these sorts of errors will often object if adults use the child's own pronunciation of the word. In fact, it has been argued that children not only perceive contrasts that they do not appear to produce, but they may even make distinctions in their productions of such contrasting words, albeit not the ones adults are expecting to hear (Braine 1976; MacKen and Barton 1980).

At first glance, another reason it seems implausible that language learners' perceptual representations of words are the cause of mispronunciations is all the research on the speech discrimination capacities of young infants. Given that young infants are so adept at detecting fine distinctions between speech sounds, why should they not have sufficiently detailed representations of their first words? Still, there are indications...
that some of the phonemic contrasts that were easily discriminated at a younger age in speech perception tasks may present problems for older infants when the contrasts occur between words. An early study by Shvachkin (1973) with Russian-learning infants suggested that they might only gradually perceive the full range of Russian phonemic contrasts in words. In a longitudinal study with infants between the ages of 10 and 24 months, Shvachkin attempted to teach children minimal word pairs. Each child was taught to associate a nonsense word with a particular toy. The sound patterns of these items differed by a single phonetic feature. The child’s task was to produce the appropriate toy from an array of different toys upon hearing the word. By the end of the study, about half of the children were responding correctly to the whole range of contrasts. However, initially the children could only respond to a few discriminations. Moreover, some discriminations appeared to be easier than others, and the children were reported to follow the same developmental sequence with respect to the kinds of contrasts that were discriminated at various ages. (In this respect, Shvachkin’s data appear to bear out Jakobson’s claims about a developmental progression in phonemic acquisition.)

Subsequently, several American researchers carried out investigations along the same lines as Shvachkin (Edwards 1974; Garnica 1973). In Edwards’s study, the infants tested were considerably older (20–47 months), but there was still evidence for differences in order of acquisition of certain contrasts. In particular, the children had more difficulty with contrasts involving fricatives than with other types of distinctions. More recently, Werker (1994) has reported on an investigation using a version of the word-learning paradigm with infants. Infants at 15 months of age were shown videos of two different nonsense objects, each of which was paired with a different nonsense syllable. During the test phase, infants heard one of the original nonsense syllables paired with both of the test objects (correctly in one case, but wrongly in the other). Although Werker found that infants at this age did react to incorrect pairings when the original test words differed in several phonemes, she did not find evidence that infants at this age learned a minimal pair distinction.

How can we explain why older infants appear to have so much difficulty discriminating the kinds of contrasts that were easily discriminated by them at 2 months of age? One possible explanation focuses on the obvious differences between the tasks that the infants are expected to perform at the two ages (Jusczyk 1977, 1985b; Locke 1988). The 2-month-old simply has to pick up an acoustic contrast between a pair of syllables, but the task for the older infant is considerably more complex because it involves associating a particular sound pattern with a particular referent. Moreover, success on this task requires representing, encoding, and recalling a particular sound pattern in sufficient detail to discriminate it from sound patterns that could be similar to it in many unforeseen respects. Much as task-demand differences between babbling sounds and producing particular sounds to convey a meaningful word may affect children’s control over their articulatory routines, so too might discrimination processes be affected by having to encode and access the right meanings in conjunction with particular sound patterns.

The child’s attentional focus is at the very least split between the visual properties of the object and its similarities and differences to other objects in the task, and the acoustic and/or phonetic properties of the speech sounds used to label the objects. It is not even inconceivable that infants at this age might attend more closely to the visual stimulation in such tasks than to the auditory stimulation. This may be true of word learning in the real world, as well as in these experimental settings. This brings up another point to bear in mind. Although, on the surface, word-learning tasks reproduce certain aspects of learning the names of objects in the real world, they may lack a critical component, namely, the child’s intrinsic motivation to learn the name of a particular object. It is possible that the latter factor does affect the encoding of words in these experimental tasks. Indeed, children tend to perform better on these types of discrimination tasks when they are tested on words that they already know, as opposed to ones that they have learned during training (e.g., Barton 1976).

In addition to the possibility that more complicated task demands affect infants’ discriminative capacities on word learning tasks, it must be noted that we know very little about the kind of information that goes into infants’ early representations of words. It has been suggested by many investigators that the early representations may only include global descriptions of a word’s sound pattern such as its prosodic structure and gross acoustic shape (Ferguson and Farwell 1975; Jusczyk 1985b, 1992; Menyuk and Menn 1979; Studdert-Kennedy 1986; Walley 1993; Water- son 1971). Although such representations are likely to include information about some salient acoustic features (e.g., manner of onset, presence of aperiodic noise, nasal resonance, etc.), this information is likely to be incomplete and not organized into phonetic segments. Because younger infants’ lexicons have fewer acoustically similar items (lexical neighbors) than do adults, they could rely on less-detailed acoustic information
to recognize words (Charles-Luce and Luce 1990, 1995; Jusczyk 1992; Walley 1993).

Thus, the fact of the matter is that there is still a great deal of uncertainty about the nature of infants' perceptual representations of words, including those that show up early in production. Nevertheless, the belief that infants' perceptual representations of speech are reasonably close approximations to phonetic descriptions led to proposals of two-lexicon models (Ingram 1976; Menn 1983; Spencer 1986; Vilman 1982). The basic premise of these models is that an input lexicon is used in recognizing words, and a separate output lexicon, containing information derived from the input, is used for word production. Among the arguments advanced in support of these models is the fact that children can perceive many more sound distinctions than they produce (Matthei 1989). Critics of two-lexicon models have pointed to the redundancy involved in storing information relevant to a word's use, such as its syntactic, semantic, and morphological information in two different places (Wheeler and Iverson 1976).

However, a reasonable alternative to the two-lexicon model is what has been termed a two-entry approach (Matthei 1989; Wheeler and Iverson 1976; cf. Menn and Matthei 1992). Such a model has only a single lexicon containing relevant syntactic, semantic, and morphological information. However, two-entry models have separate access routes to the information contained in the central store for perception and production. Thus, one means of accounting for some of the differences noted in the development of phonological categories in perception and production is to view them in terms of a two-entry approach to the lexicon. Each type of access route might develop independently, at least initially. Furthermore, given the differences in the sensory and motor systems underlying perception and production, respectively, it would not be unreasonable to expect some sort of developmental décalage in terms of when the learner gains command of a word in perception and production. The fact that the vocal tract undergoes large changes in shape and dimensions, plus the amount of coordination required to sequence the articulators properly, might well cause productive capacities to lag behind perceptual ones.

Assessing the Relation between Changes in Perception and Production

One implication of a two-entry approach to the lexicon is that perception and production develop relatively independently of each other. Their main point of contact comes through links to a common meaning in the lexicon. This means that perceptual and productive processes are more likely to influence each other after links to common meanings have been established. However, does this imply that there are no links between perception and production prior to this point? Ruling out any mutual influences at younger ages appears to be too strong a claim in view of findings indicating that 5-month-olds show some ability to match vowel sounds to the sight of appropriate mouth movements (Kuhl and Meltzoff 1982, 1984; MacKain et al. 1983). Furthermore, infants' own productions are also an available source of perceptual input. Certainly, one would expect that some play in producing sounds in babbling involves exploring the acoustic consequences of various articulatory gestures. The point here is that as "meaning" and the desire to be understood become the child's primary preoccupations, there is greater pressure to bring productive representations of words more closely in line with perceptual representations.

Maybe we should step back for a moment and ask when the lexicon actually begins to develop. Some of the evidence suggests that infants may begin storing information about sound patterns at some point during the second trimester of their first year (Hallé and Boysson-Bardies 1994a; Hohne et al. 1994; Mandel et al. 1995). Moreover, there is evidence indicating that infants begin to comprehend a few words at some point between 8 and 10 months of age (Benedict 1979; Huttenlocher 1974). Thus, there is some reason to believe that the formation of the lexicon occurs during the latter half of the first year. Consequently, one might expect to see the perception and production systems beginning to interact during this period. Let us consider what kinds of changes are taking place in perception and production at this time and the degree to which they may be related.

With respect to perception, the latter half of the first year is characterized by a marked increase in sensitivity to the structure of native-language sound patterns. This shows up with respect to phonetic, phonotactic, and prosodic features. The infant displays a sensitivity to the frequency with which these kinds of patterns appear in the input. Moreover, infants advance in their abilities to segment words from fluent speech and to detect the natural groupings of words that the language imposes on utterances (i.e., clauses and phrases). With respect to production during this same period, most of the infants' vocalizations still consist of babbling. However, it is at this point that native-language input begins to influence the
kinds of strings that are produced in babbling. As the work of Boysson-Bardies and her colleagues shows, the kinds of vocalic and consonantal elements that appear in babbling are influenced by the distribution of such elements in the input that the infant is receiving. Similar changes are evident in the prosodic structure of babbling, as the work of Levitt and her colleagues also indicates. Moreover, there are even cross-linguistic differences in the frequency with which certain kinds of syllable structures appear in babbling that begin to show up late in the first year (Vihman 1993b). What these findings demonstrate, then, is that linguistic input affects infants’ productions, just as it also affects their perceptual capacities. However, since the input comes through the perceptual system, these examples are largely demonstrations of how perception affects changes in production that occur during this time.

The influences of production on perception are not as obvious. However, Vihman has mentioned several interesting possibilities. One of these is her notion of an articulatory filter that affects what learners attend to in speech input. More specifically, children’s abilities to produce some kinds of sounds better than others could affect their attention and memory for words by favoring words that include sounds that they have mastered. Although Vihman herself sees the articulatory filter as primarily affecting word production rather than comprehension, it would be interesting to explore the extent to which infants’ recognition of words might show these same influences. For example, if an articulatory filter serves to focus attention on certain patterns in the input, then this might increase the chances that words embodying such patterns are the most likely to be added to the lexicon. Hence, one might find that the proportion of such items in the lexicon is greater than otherwise might be expected.

A second way the child’s productive schemes could influence perceptual processes is in the decline of sensitivity to nonnative contrasts. Specifically, Vihman (1991; see also Studdert-Kennedy 1991a) indicates that when these declines appear, infants have already begun to favor in their productions the kinds of articulatory gestures that are used in the native language. She points out that because infants are in the process of matching patterns that they hear to their developing vocal motor schemes, this may mean that nonnative sounds are not “meaningfully processed at this point.” She suggests further that perhaps the same thing occurs for native-language sounds that are not within the child’s articulatory repertoire at this stage. This last point is intriguing, in light of the difficulties that infants have in word-learning experiments—although it is by no means clear that their difficulties are confined to words that include segments that they do not yet produce.

In considering how changes in perception might affect production and vice versa, we should not lose sight of the fact that some determinants of developmental change in each of these systems will have little to do with what happens in the other system. Thus, there are changes in production that will have to do with maturation and a greater ability to coordinate one’s own motor patterns. Similarly, there are changes in speech perception that are more likely to reflect changes in memory and attentional capacities than any influences from speech production. Each system brings with it a particular set of constraints that affects performance. For example, from the point of view of ease of operation, there are pressures on the developing articulatory system to produce all sounds in much the same way. However, in this case the resulting sounds would be difficult to discriminate. From the point of view of perception, it is ideal to have the sounds that one uses be as distinct as possible so that they will be less confusable. However, this would increase the range of different kinds of articulatory gestures that have to be mastered and coordinated. What happens in the sound systems of natural languages is some sort of compromise between these two extremes (although there are very likely other factors that also affect how the sound system of any given language is structured, as Anderson 1981 has suggested). Each language has its own solution to this problem.

Lindblom and his colleagues (e.g., Lindblom 1986, 1992; Lindblom, MacNeilage, and Studdert-Kennedy 1983) have proposed a model of how the interaction between perceptual and articulatory constraints may lead the language learner to develop a phonological system from more-global lexical representations. They identify discriminability and pronounceability (a tendency to reduce articulatory complexity) as important constraints on listening and speaking. They also propose that a combination of production and perception constraints, which they call sufficient contrast, helps to shape phonetic inventories in languages. Their assumption that phonological units and rules emerge as a consequence of lexical development is similar to the assumption that underlies the view presented below.

There is another way interactions between the developing perception and production systems may affect the growth and character of language
learners' knowledge of native-language sound patterns. Pressures to coordinate the way that these systems function and to relate the perceptual representations of words to the articulatory representations used to produce these words may force learners to derive a more abstract description that captures relevant generalizations that apply to both systems. The abstract description in this case is phonology. Thus, one suggestion is that it is the coordination of perceptual and productive representations that may lead the language learner from a more global representation of sound patterns of words to one that is structured with respect to phonetic segments.

The notion that phonological descriptions emerge from efforts to coordinate the outputs of perception and production systems during development is not new. Bever (1975, 1981) proposed something along these lines in his reflections about the purpose of a grammar during development. Speaking more generally about grammar than about phonology per se, he noted that the reason that a psychogrammar exists is because of the vital role it plays during language acquisition, much of which occurs during the first five years of life. The psychogrammar is needed during that period to mediate between the systems of speech production and perception. It is the internal translator that regulates conflicting capacities which arise as each of the two systems of speech develop separately: if one system gets ahead of the other the psychogrammar can equilibrate their capacities. (Bever 1975, 65)

Thus, according to Bever, the whole reason for the existence of a grammar is that it helps to coordinate the systems of perception and production during development. He also argued that once this coordination was achieved, the grammar need not play any direct role in on-line processing (i.e., grammatical rules are not necessarily executed as steps during processing, nor does processing require computing the kinds of successive representations that are associated with derivations of sentences). Also, he speculated that the successful completion of the coordination process was what effectively shut down the critical period for language learning. In his view, this occurs earlier for the phonological system than it does for either syntax or semantics.

The phonological structure of a grammar as a mediating system becomes unused earlier than the semantic or syntactic structures. This is not to say that the phonological system is easier to learn—only that sufficient data for its complete acquisition are available to the child at an early age: once the speaking and identifying capacities are equilibrated, further phonological learning stops, and the constructed mapping lies fallow. . . . The critical age for phonological learning will be younger than for syntax and semantics. (Bever 1981, 195)

The framework within which Bever proposed his explanation was the Language Acquisition Device, which was part of the innate endowment of the language learner. Among other things, the LAD was expected to include information about formal and substantive linguistic universals (see chapter 2). Thus, with respect to the phonological component of the LAD, the infant would use knowledge of formal and substantive universals to find a successful way of resolving the outputs of both the perception and production systems.

I am a bit more agnostic regarding the amount of innate knowledge that language learners have regarding linguistic universals. Thus, I would like to suggest that some generalizations about phonological universals may actually arise from the process of coordinating the outputs of the perception and production systems. What may be innate about the whole thing is not so much a set of categories or constraints on formal rules but rather the drive to find a more general abstract system and a means of representation that coordinates the outputs from the perception and production systems. Thus, the coordination between the perception and production systems need not depend on the existence of a hardwired device such as the LAD.

Finally, the notion that production and perception may become coordinated and more integrated during the course of development seems to fit with what Gathercole and Baddeley (1993) have noted about the component of working memory that they call the phonological loop. In their view, the phonological loop is a system that is specialized for the storage of verbal material. It has two subcomponents: (1) the phonological store, which represents material in a phonological code that decays over time; and (2) an articulatory rehearsal process, which refreshes and maintains the decaying items in the phonological store. Spoken information gains direct access to the phonological store without articulatory rehearsal. Consequently, it is possible that this store is active in perception even during the first year. Although Gathercole and Baddeley indicate that the phonological loop is present and functioning from the preschool years onwards, they also note that there is little evidence that the articulatory rehearsal process is fully operative at this stage. For example, articulatory suppression, which prevents the rehearsal process, does not have the same