How Attention to Sound Properties May Facilitate Learning Other Elements of Linguistic Organization

It seems pretty clear that the basic speech perception capacities that infants possess should prove useful in learning how the sound structure of the native language is organized. What is less obvious is that these same perceptual capacities could play a role in acquiring information about other levels of language structure. Languages vary in terms of their sounds and their meanings, and the way that sound patterns are organized does not bear an obvious relation to how patterns of meanings are organized in the language. The simple notion that similarities in sound patterns predict similarities in word meanings is easily disproved by considering a few examples in any language. Thus, in English, the sound patterns of the words "rat" and "cat" are more similar to each other than either is to the word "mouse," yet a consideration of the referents of each of these items would clearly find "rat" and "mouse" are more closely grouped together than either is to "cat." Similarly, "mushrooms" and "mush" are less semantically related than "mush" and "cereal," and so on. Consequently, any approach that looks for a very direct mapping between similarities at the level of sounds and the level of meanings is doomed to failure.

However, there are other ways the sound structure of native-language utterances could bear a relation to other levels of linguistic organization. One possibility, often discussed, is that elements of the syntactic organization are marked in the acoustic structure of utterances. As was noted in discussing McNeill's work, this possibility was considered and rejected about thirty years ago on the grounds that units derived from the acoustic signal do not necessarily correspond to the critical units in the linguistic analysis of the utterances. For example, pauses in speech can be the result of hesitations that have to do with finding the correct word as well as for syntactic reasons. In fact, McNeill believed that correctly interpreting many of the kinds of acoustic changes that occur in fluent speech could only come about if one already was in possession of a grammar. Without a grammar, learners could only track the physical changes in the signal, not the perceived contour that is correlated with the syntactic organization. Hence, the idea that the child could learn the syntax from what happens to the speech signal seemed unlikely to him.

Actually, McNeill's arguments were directed at whether there was sufficient information in the speech signal to pull out the syntactic structure. He was reacting to a very strong version of what is now usually called prosodic bootstrapping. More specifically, he was evaluating the possibility that one could recover the entire syntactic description of an utterance from the speech signal. Thus, his arguments were directed at this strong view. He never did claim that there was no useful information about syntactic organization to be derived from the acoustic signal.

It was not until the 1980s that researchers in language acquisition seriously considered how information in the signal could facilitate the acquisition of syntax. Until then, researchers more or less sidestepped the issue of how the child ends up with the right pieces of the speech signal to use for syntactic analysis. For most accounts, the starting point was that the child was already dealing with clauses and simply had to discover their internal syntactic organization. How the child succeeded in isolating the right units, and not fragments from different units, was not seriously considered until investigators such as Gleitman and Wanner (1982; Gleitman et al. 1988) and Peters (1983, 1985) began to suggest that information in the speech signal itself could be used for this purpose. The basic notion was that prosodic (and other) markers in the speech signal provide clues to syntactic units such as clauses and phrases. In addition, there was speculation that such marking could even be helpful in working out the syntactic tree (i.e., the hierarchical arrangement of grammatical units) underlying the utterance (Gleitman et al. 1988).

One reason for the renewed interest in the information available in the speech signal had to do with a reaction to proposals based on learnability accounts of language acquisition, such as those offered by Wexler and Culicover (1980). In particular, these accounts appeared to demand exposure to a range of very complex sentence structures in order to select the grammar that correctly characterizes utterances in the native language that the learner is acquiring and to rule out possible alternative grammars. Morgan (1986) demonstrated that access to information in the speech signal that helps to bracket syntactic units appropriately could effectively reduce the degree of complexity required for learners to select the grammar that corresponds to the native language that they are acquiring.

It is important to note that information available in the signal was seen as a means of complementing, but not supplanting, other sources of information that learners draw upon to discover the syntactic organization of their language. For example, although there might be information in the signal to help

in correctly bracketing the input, the learner still needed some means of labeling the brackets appropriately (i.e., assigning each such unit to the correct grammatical category) and of working out the relations among the items clustered within a particular bracket. Hence, prosodic bootstrapping accounts were intended to show how information in the speech signal could interact with innate linguistic capacities or with other cues to sentence structure and lead to the acquisition of a native language (Morgan 1990).

The notion that learners may rely on information in the speech signal to help in the discovery of syntactic organization has come to be known as "prosodic bootstrapping." Still, the speech signal contains other potentially valuable sources of cues that are not necessarily prosodic in nature. For example, some languages use *concord morphology,* wherein words that appear with certain syntactic units share similar affixes. Also, the recurrence of function words in certain sentential positions could be helpful in signaling certain types of syntactic units. One indication that all of these kinds of cues are potentially helpful in learning the syntactic organization of a language conies from a study of adults learning artificial grammars (Morgan, Meier, and Newport 1987). In separate experiments, Morgan and colleagues found that the presence of prosodic markers, concord morphology, or function words resulted in faster and more accurate learning of artificial grammars. It is conceivable that some of the success that the adults had with these cues is attributable to their knowing a first language and how comparable cues operate in it. Nevertheless, the results of this study offer an important demonstration that the presence of these kinds of cues can facilitate the learning of syntactic relations.

What conditions must hold to seriously entertain the proposition that prosodic bootstrapping occurs in first-language acquisition? First, there must be some indication that there are acoustic correlates of syntactic organization present in speech. Second, the potential prosodic correlates must be ones that infants are able to detect in speech. Third, there must be some indication that infants actually rely on these correlates in organizing the input. Let us first consider the evidence for these conditions. Then, we will consider how other possible sources of information in the signal that are not prosodic in nature could also bear on the acquisition of the syntax of a native language.

Evidence for Prosodic Marking of Grammatical Units

There are a number of cues associated with prosody, such as fundamental-frequency changes (i.e., intonation contours), stress patterns, pausing, and durational differences that could serve as potential markers of units in the speech stream. It has been noted that clause boundaries in English are often marked by pauses, increases in the duration of syllables preceding the boundary, and changes in fundamental frequency (e.g., Klatt 1975; Luce and Charles-Luce 1983; Nakatani and Dukes 1977). For example, Cooper and Paccia-Cooper (1980; see also Martin 1970) reported evidence of longer pauses at clause boundaries. Similarly, a number of investigators have found evidence of segmental lengthening in the syllable immediately preceding syntactic boundaries (Cooper and Paccia-Cooper 1980; Klatt 1975, 1976; Price et al. 1991; Wightman et al. 1992). Price et al. (1991) also found evidence of pitch changes in the vicinity of syntactic boundaries. In addition, linguistic analyses of different languages point to a tendency for intonation groups to correspond to major syntactic units such as clauses or their important components (Bolinger 1978; Cruttenden 1986; Selkirk 1984).

There are also indications that some of these potential cues to syntactic boundaries do play a role *in* how listeners perceive these fluent utterances. A number of relevant demonstrations have used sentences with potentially ambiguous phrasal boundaries (Collier and t'Hart 1975; Lehiste, Olive, and Streeter 1976; Scott 1982; Scott and Cutler 1984). In one such study, Streeter (1978) independently manipulated intensity, duration, and intonation cues. She found that each of these cues could affect the way that ambiguous sentences were interpreted. More recently, Price et al. (1991) conducted an extensive investigation using a systematic set of materials that were produced by radio announcers on different occasions. In addition to analyzing these materials for acoustic cues to boundaries, investigators collected perceptual data from adult listeners that confirmed that changes in syllable durations, fundamental frequency, and pausing were used as indices of syntactic units. Finally, the results of several investigations suggest that listeners are sensitive to speech cues that are correlated with phrasal units in unfamiliar foreign languages (Pilon 1981; Wakefield, Doughtie, and Yom 1974).

Nevertheless, it must be noted that many of the same acoustic changes that frequently coincide with important syntactic units in speech also occur in utterances for nonsyntactic reasons (Beckman and Edwards 1990; Grosjean and Gee 1987; Nespor and Vogel 1986; Vassiere 1981). For example, lengthening may occur to distinguish a voiced stop consonant from a voiceless one (Klatt 1976).

Intonation may convey stylistic and affective attributes of the talker (Fairbanks and Pronovost 1939; Lieberman 1961; Williams and Stevens 1972). Consequently, if listeners were to rely on any one of these cues for information about grammatical units, they still would need some other mechanism to let them know when the cues were actually relevant to syntactic matters. Hence, we cannot assume that listeners can read the entire syntactic organization of a sentence from an analysis of the speech signal.

The upshot of all this is that although syntax is an important factor, it is not the sole determinant of the organization of suprasegmental information in a sentence. At the same time, the fact that a particular prosodic cue can serve other functions does not mean that it is useless for signaling important grammatical units. First, as the data reviewed above indicate, these cues have been shown to affect the interpretation of sentences. Second, prosodic cues to phrasal boundaries commonly act in combination with one another. For example, pitch changes at the ends of clauses in English tend to occur in combination with increased final syllable durations and pausing (Price et al. 1991). As with other cases involving multiple cues, the contribution of any one cue may not be as important as the summed tendency or possible interaction of several different cues. Thus, it is the constellation of cues from the speech signal, rather than any single cue, that is likely to provide information helpful to the grammatical organization of utterances.

So far, the evidence considered only relates to the prospect that there are potential prosodic markers of syntactic units in fluent speech between adults. The acoustic characteristics of child-directed speech are known to differ from adult-directed speech in several ways (Ferguson 1977; Garnica 1977; Grewel 1959). Many prosodic features have been shown to be exaggerated in speech that is addressed to children. Among other things, the average pitch of child-directed speech is known to be higher, the pitch range more exaggerated, and the durations of content words longer than for adult-directed speech (Garnica 1977). Moreover, these kinds of prosodic differences between child- and adult-directed speech ha been observed across a range of different languages (Fernald and Simon 1984; Fernald, Taeschner, Dunn, Papousek, Boysson-Bardies, and Fukui 1989; Grieser and Kuhl 1988; Papousek, Papousek, and Haekel 1987. Stern, Spieker, Barnett, and MacKain 1983). In addition, more consistent use of pitch changes at clause boundaries has been observed for child-directed than for adult-directed speech (Garnica 1977; Stern et al. 1983) Similarly, Broen (1972) reported that pauses were used more often at clause boundaries, and Bernstein Ratner (1986) found more pronounced segmental lengthening in the vicinity of these boundaries.

Other recent investigations have provided strong evidence for the prosodic marking of clause boundaries in child-directed speech, although the picture with respect to the marking of sub-clausal units (such as syntactic phrases) is more mixed (Fisher and Tokura 1996; Jusczyk et al. 1992; Lederer and Kelly 1991). For instance, Fisher and Tokura (1996) found no consistent prosodic marking for units smaller than clauses in the sample that they analyzed, whereas Jusczyk et al. (1992a) found correlations of pitch and duration cues with subject/predicate phrase boundaries in their samples. Lederer and Kelly (1991) reported that even minor differences in syntactic structure (e.g., prepositional phrases) were marked in the samples that they analyzed. We will return to this issue when we consider results pertaining to infant's sensitivity to phrasal boundaries.

There are indications that prosodic changes frequently coincide with syntactic boundaries in speech between adults. Changes in pitch, syllabic duration, and pausing have been shown, both individually and in combination, to affect the perception of syntactic boundaries. In addition, many of these same prosodic cues appear in an exaggerated form in speech directed to children.

Sensitivity to Prosodic Markers in the Input

Although it is interesting that some prosodic marking of syntactic units is present in the input received, we need to know whether infants are actually sensitive to its occurrence. For example, there is information to be perceived in the ultraviolet region of the spectrum but, unlike bees; humans are not capable of seeing it. Likewise, in order for prosodic cues to play any explanatory role in how the syntactic organization of language is acquired, we must show that infants are capable of responding to this information. Thus, the first order of business is to devise a means to find this out.

Hirsh-Pasek et al. (1987) reasoned that if infants respond to prosodic markers in the input, then they should prefer to hear speech that is segmented in accordance with these markers as opposed to speech that is segmented inappropriately. Hence, they collected samples of speech from a young woman who was talking to a 19-month-old. They excised a set of passages that were five to seven sentences in length, and they inserted a series of 1-second pauses in each of them. Two versions of each passage

were produced. In one case, all the pauses were inserted at the boundaries between two clauses (we will refer to these as the *coincident versions);* in the other case, an equal number of pauses were inserted but between two words in the middle of a clause (we will call these the *non-coincident versions).* It is important to note that the location of the pauses in the coincident versions coincided with other markers of the clause boundary such as syllable lengthening and pitch declination. Hence, for the coincident versions, the pause locations cooperated with whatever other prosodic markers of clause boundaries were present, whereas for the non-coincident versions, the pause locations competed with other potential markers of clause boundaries.

Hirsh-Pasek and colleagues hypothesized that if infants are sensitive to the prosodic marking of clauses in the input, they would prefer to listen to the coincident versions of their samples than to the non-coincident versions.[1] In their first experiment, they used the head-turn preference procedure to present samples of both types to 10-month-olds. The infants listened significantly longer to the coincident than to the non-coincident versions of the samples. In a second experiment, they demonstrated that 7-month-olds also displayed the same listening preferences for the coincident versions. Hirsh-Pasek and co-workers interpreted these findings as an indication that infants as young as 7 months old are sensitive to the presence of prosodic markers to clausal units. A subsequent investigation by these same researchers examined whether the exaggerated prosody of child-directed speech may have enhanced the infants' ability to detect markers to clausal units (Kemler Nelson et al. 1989). The same woman who had recorded the child-directed speech in the previous experiment produced adult-directed speech samples for the new study. Once again, passages between five and seven clauses in length were chosen, and two versions of each one were prepared by inserting 1-second pauses either at clause boundaries (coincident versions) or in the middle of clauses (noncoincident versions). As in the previous study, the 8,5-month-olds who heard the child-directed speech displayed a significant preference for the coincident versions. However, the infants tested on the adult-directed speech samples showed no such preference, prompting Kemler Nelson and colleagues to conclude that the exaggerated prosody of the child-directed speech samples may have facilitated the detection of markers of clausal units in these samples.

A subsequent investigation by Morgan, Swingley, and Miritai (1993) provided converging evidence that English-learning infants are sensitive to the prosodic marking of clause boundaries in the input. In particular, they used a paradigm in which infants had to respond by turning their heads whenever a noise stimulus occurred. The infants were better able to detect the noise when it occurred between words from different clauses than between words within the same clause. The suggestion here, as in click-detection experiments with adults (Abrams and Bever 1969), is that the noises are more easily detected when they do not interrupt perceptual processing units (in this case, clauses).

How can we be certain that infants in these studies were actually responding to the prosodic markers rather than to some other information available in the utterances? One demonstration that prosodic cues are important is to show that they are sufficient to produce the longer listening times to the kinds of coincident samples that Hirsh-Pasek et al. (1987) used in their study. To explore this possibility, Jusczyk (1989) low-pass filtered Hirsh-Pasek et al.'s coincident and noncoincident samples at 400 Hz to eliminate most of the available phonetic information in the samples. Six-month-olds tested on these low-pass-filtered samples displayed the same listening preferences for the coincident versions. Hence, even when most of the phonetic information was removed from the signal, infants still responded to the way the pauses were related to the prosodic information in the utterances. Therefore, these findings are consistent with the view that infants are sensitive to the type of prosodic marking that could ultimately provide cues about clause boundaries in utterances.

What is the basis of infants' sensitivity to prosodic marking of clausal units? Is this sensitivity specific to native-language patterns that the infants have experienced, or is there a more general basis for the way the infants respond? There are indications that the kinds of prosodic changes that occur at clause boundaries in English also occur in many other languages (Cruttenden 1986). Hence, one possibility is that infants are responding preferentially to the coincident versions of the samples for reasons other than the fact that they have picked up something specific to English prosodic structure. Rather, their responsiveness to these types of changes may reflect a more general bias for processing auditory input. What kind of evidence might lead to such a conclusion? A demonstration that infants are sensitive to the marking of clause boundaries in a language, regardless of any experience with it, might be an indication of a more general bias. Also, evidence that the cues that are important in signaling clause boundaries are the same as those used to signal other kinds of event boundaries in auditory perception

would seem to favor a more general mechanism, rather than one that is specific to language.

There have been several investigations of how infants respond to clause boundaries in other dialects of their own language and in unfamiliar languages. In one study that is reported in Polka, Jusczyk, and Rvachew (1995), American 9-month-olds were presented with samples produced by a female talker of British English. As in the Hirsh-Pasek et al. (1987) study, the samples included pauses either at clause boundaries or between words in the middle of clauses. Even though the speech patterns of British English were unfamiliar to them, the infants still listened significantly longer to the coincident versions of the samples. In another investigation, Jusczyk (1989) reported that American 4,5month-olds not only listened significantly longer to coincident versions of samples in English, but they also showed the same pattern of responding for utterances in an unfamiliar language, Polish. Interestingly enough, by 6 months of age, the American infants no longer showed a preference for coincident over noncoincident versions of the Polish samples, even when these were low-pass filtered to remove potentially distracting phonetic information. One possible interpretation of these results is that a language-general ability to perceive prosodic marking of clausal units by the 4,5-month-olds has given way to more language-specific processing of the utterances, and that by 6 months, the rhythmic patterns of Polish mark it as normative input.

To this point, the findings are consistent with the notion that, at least initially, there is a general, rather than language-specific, basis for sensitivity to prosodic marking of clausal units in utterances. However, this picture is complicated by the results of a study that investigated American 4,5-month-olds' perception of Japanese utterances. Mandel, Jusczyk, and Mazuka (1992) presented American infants with child-directed speech samples produced by a Japanese mother. They found no evidence that infants at this age listened significantly longer to coincident versions of these samples than they did to noncoincident versions. Another investigation using new samples produced by a different Japanese mother yielded the same pattern of results, that is, American 4,5-month-olds| did not show a significant listening preference for coincident versions 1 (Jusczyk, Mazuka, et al. 1993).

While it is always difficult to interpret null results, the behavior of American 4,5-month-olds toward the Japanese samples certainly seems to contrast with how they responded to the Polish samples. Two possible explanations for the discrepancy come to mind. The first is that the basis for the detection of prosodic markers to clause boundaries is basically language-specific. By this line of reasoning, at least initially, Polish patterns (but not Japanese ones) are simply close enough to the English ones to have engaged the processing routines that American 4,5month-olds use for their native-language utterances. The second possible reason is that the use of such prosodic markers is somehow tied to the rhythmic properties of languages. Japanese is organized around the mora as a rhythmic unit, Polish and English are not. Perhaps this basic feature of rhythmic organization is one that even 4,5-month-olds pick up. There are indications that although French newborns are sensitive to changes in numbers of syllables, they do not react to changes in the numbers of morae present in utterances (Bertoncini 1993; Mehler et al. 1995). Thus, it may be that the fact that rhythmic patterns of Japanese are perceived to be normative by American infants earlier than the rhythmic patterns in Polish. If so, then this may cause the American infants not to process prosodic cues in the same way as they would for utterances with a rhythmic basis similar to that of English. Of these two possible explanations, the first seems less complicated and hence more plausible. However, further research with other languages is needed to decide between them or among other possible alternatives.

At the same time, there are hints of possible nonspeech parallels to the sensitivity that infants show to the prosodic marking of clausal units in utterances. Specifically, studies with musical stimuli (Mozart minuets) indicate that 4,5month-olds listen significantly longer to samples with pauses inserted at musical phrase boundaries than they do to samples with pauses inserted in the middle of musical phrases (Jusczyk and Krumhansl 1993; Krumhansl and Jusczyk 1990). Interestingly enough, the cues that appear to signal musical phrase boundaries for infants are a decline in pitch and a lengthening of the final note at the musical phrase boundary. These cues parallel ones associated with clause boundaries (i.e., decline in pitch and clause-final syllable lengthening). These parallels observed between the perception of phrase boundaries in music and clause boundaries in speech are tantalizing. They indicate that we cannot totally preclude the possibility that common mechanisms underlie the perception of event boundaries for both speech and music.

Although it is certainly valuable for language learners to have some means of locating clausal units in the input, it would be even more valuable for them to have some way of further segmenting the input into subclausal units, such as phrases. Indeed, as Pinker (1984) has commented the ability to track the

distribution of phrases and of constituents within phrases would take the child a long way toward constructing a grammar. Obviously, prosodic marking of phrasal units is more likely for languages in which word order is more constrained (and hence, likely to keep words within the same constituent together) than for ones in which word order is free. For this reason, the means by which languages mark subclausal units in the input could vary considerably. Concord morphology and other such devices may be a more useful means than prosodic marking to identify elements of the same phrasal constituent in languages with relatively unrestricted word order. Consequently, not unlike the situation of learning about the phonetics and phonotactics of sound patterns in one's native language, learners may have to discover the specific means by which phrasal units are marked in their native language.

In a language like English, which relies on word order to signal syntactic relations, it is not unreasonable to expect that some prosodic marking of phrasal units could occur. Indeed, as noted earlier, there are indications that English-speaking adults are sensitive to potential acoustic correlates of phrase boundaries (e.g., Lehiste et al. 1976; Price et al. 1991; Scott 1982). Among the various kinds of phrasal units in English, subject phrases and predicate phrases are ones that might receive significant marking in the linguistic input. For example, sentence subjects in English function in a number of important ways: (1) they carry nominative case marking; (2) they control agreement in person and number with the verb; (3) they are usually identified as the agent of a transitive verb; and (4) they function as the topic of the sentence (Givon 1979). Similarly, predicate-verb phrases define the architecture of the sentence. The verb phrase, particularly the verb, is responsible for assigning thematic roles in the sentence. In fact, the learning of verbs and verb-phrase structure is seen as pivotal to the induction of grammar in some current theories of language acquisition (e.g., Fisher et al. 1994; Gleitman 1990, 1994; Golinkoff et al. 1987; Naigles and Kako 1993; Pinker 1989). Finally, there is evidence that, at least under some circumstances, intonation groups do align with the subject-predicate division in sentences (Beckman and Edwards 1990; Cruttenden 1986).

For the reasons just described, Jusczyk et al. (1992a) decided to investigate English-learning infants' detection of prosodic marking of subject and predicate phrasal units in utterances. In a series of experiments, using materials drawn from either child-directed spontaneous speech or from stories read to a child, they inserted pauses either at boundaries between subject and predicate phrases (coincident versions) or at locations in the middle of phrases (noncoincident versions). Several interesting findings emerged from this investigation. First, 9-month-olds, but not 6-month-olds, proved to be sensitive to the location of the pauses. That is, only at 9 months of age did infants listen significantly longer to coincident versions than they did to the noncoincident versions. This was confirmed both in cross-sectional comparisons involving different infants at the two ages and in longitudinal comparisons at 6 and 9 months of age with the same infants. The finding that 6-month-olds do not respond to marking of phrasal units is interesting in light of the findings reviewed above that indicate English-learning infants as young as 4,5-months of age are sensitive to the marking of clausal units. This may be an indication that infants require more extensive experience with their native language before they detect prosodic markers of phrasal units in the input.

The case that infants in Jusczyk and coworkers' studies were responding to prosodic markers in the passages is strengthened by the results of additional experiments with low-pass-filtered versions of the stimuli. Once again, the 9-month-olds listened significantly longer to coincident than to noncoincident versions of the samples, even when they were low-pass filtered. Furthermore, acoustic analyses of the passages indicated that relative to the noncoincident versions, the coincident versions were more apt to have syllable-final lengthening and pitch drops just prior to where the pauses were inserted (i.e., at the phrasal boundaries in the coincident versions).

At first glance, Jusczyk and his colleagues' findings seem to offer some hope for proponents of the view that, at least for languages like English, infants may be able to discover the syntactic organization of utterances (including the hierarchical ordering of different phrases) directly from prosodic features (e.g., see Lederer and Kelly 1991 for a suggestion along these lines). However, as mentioned above, the likelihood is that prosodic marking of phrasal units would not be equally effective for all languages. Moreover, even for a language like English, the situation is much less straightforward than it first appears. As has been observed in many contemporary accounts of prosodic phonology, prosodic boundaries do not always map directly onto syntactic ones (Hayes 1989; Nespor and Vogel 1986; Selkirk 1981). Indeed, mismatches in the prosodic and syntactic organization may occur even in the simple sentences that are directed to infants acquiring language. Consider the following two sentences.

(1) Mary ate the cake.

(2) She ate the cake.

In (1), the talker is likely to produce prosodic boundary cues after the subject NP, "Mary." However, in (2), even two-year-old talkers (Gerken 1991; Gerken 1994b) either produce no prosodic boundary cues or produce them between the verb and the object NP, "the cake." This is because of the prosodic structure of English. The elementary unit of rhythmic organization is the "foot." In English a foot consists of a strong syllable followed by either zero or one weak syllable (Gerken 1994a, in press; Hayes 1982; Selkirk 1980). Because a weakly stressed pronoun .subject cannot begin a foot in (2), it is said to be unfooted (Gerken 1994a, in press). However, at the level of a phonological phrase, it will be joined into the same unit as the following stressed verb-that is, the subject and verb form a prosodic unit. Hence, there is no prosodic marking of the syntactic boundary between the subject and the predicate phrases in (2). By comparison, Mary has a strong first syllable in (1), and therefore is in a foot separate from the following verb. Hence, in this case, a prosodic break is possible between the subject and the verb.

The point here is that in cases with a pronoun subject, like (2), the learner who is looking for prosodic clues about the internal syntactic constituents of the utterance either would receive no information or, maybe worse, would get misleading information about what the major constituents are. It turns out that only a small percentage of the spontaneous speech samples (about 15 percent) used by Jusczyk et al. (1992a) contained potential mismatches of the sort found in (2). Hence, most of the utterances that they used were ones in which prosodic phrase boundaries coincided with syntactic ones. (This fact may be responsible for some of the discrepancies between the results of Jusczyk and coworkers' acoustic analyses and those of Fisher and Tokura (1996) that were mentioned in the previous section.)

Gerken, Jusczyk, and Mandel (1994) explored how English-learning infants respond to utterances in which prosodic and syntactic boundaries mismatch. For this purpose, they created new materials to compare infants' responses to sentences with lexical NP subjects, as in (1), to sentences with pronoun subjects, as in (2). An example of a lexical NP type of sample is:

(3) This is a story about a little boy named Sammy. Sammy is a baseball player. Sammy can run fast. And Sammy never misses a ball. Every Saturday, Sammy plays baseball in the park.

A comparable example of a sample with pronoun NPs (after the lead-in sentence) is:

(4) This is a story about a little boy named Sammy. He is a baseball player. He can run fast. And he never misses a ball. Every Saturday, he plays baseball in the park.

Coincident versions of both types of passages were prepared by inserting a 1-second pause in all sentences (after the lead-in sentence) between the subject and predicate phrases. Noncoincident versions had the same number of pauses but these were inserted between the verb and its complement. Nine-month-olds exposed to the sentences with lexical NP subjects behaved exactly like the 9-month-olds in the Jusczyk et al. study-namely, they listened significantly longer to samples in which pauses were inserted between the subject and verb phrases than to ones in which pauses were inserted between the verb and object NP phrases. In contrast, infants who heard the sentences with pronoun subjects did not show a significant preference for either type of segmentation. This is one indication that it is prosodic phrase boundaries, and not necessarily syntactic phrase boundaries, that infants are responding to.

Of course, the English-learning child eventually has to be able to extract information about the internal organization of sentences with pronoun subjects. How could they begin to do this, given the apparent absence of prosodic marking of the subject-predicate boundary in such sentences? One possibility is that they use other (nonprosodic) cues to work out the syntactic relations of such sentences. Another possibility is that learners may benefit by contrasting cases in which pronouns and verbs belong to the same prosodic group with those in which they belong to different groups. In other words, they make cross-sentential comparisons of different utterances. Morgan, Meier, and Newport (1989) found that such cross-sentential comparisons benefited adult learners of artificial grammars. Conceivably, this might hold as well for first-language learners. Resolving the conflict in prosodic organization across cases in which a prosodic boundary sometimes occurs before a particular lexical item and sometimes after it could actually lead the learner toward the discovery of the syntactic organization. Naturally, such a solution assumes that the learner has access to such contrasting cases in the input and is able to

recognize the similarities that exist across these (e.g., the presence of familiar words or word sequences).

In fact, there are indications that the needed contrasting cases are readily available in the input that the learner receives. In sentences that involve yes-no questions, there is a tendency for the pronoun and auxiliary to form a prosodic group that may be separate from the verb. For example, a talker tends to impose a prosodic boundary just before the verb in an utterance like "Did she / throw the ball?" In these cases, the prosodic input favors a marking of the boundary between the pronoun and the main verb. But are infants sensitive to these markers? In the study that we have been considering, Gerken and colleagues conducted an additional experiment to test this possibility. They constructed passages with sentences involving inversions between a pronoun and an auxiliary, that is, yes-no questions. Pauses were inserted either before or after the main verb in these sentences to create the coincident and noncoincident versions. Nine-month-olds listened significantly longer to versions in which the pauses occurred between the subject and verb phrases (i.e., the coincident versions). The implication of these findings is that 9-month-olds are sensitive to the prosodic break between the pronoun and main verb in such sentences. So, in some situations, they may be able to assign a phrase with a pronoun subject to a separate prosodic group than the one that includes the predicate phrase. This suggests that infants at this age are, at least, positioned to notice contrasting cases in which pronouns and verbs belong sometimes to the same prosodic groups and other times to different ones.

Evidence that Prosodic Organization is Actually Used in Infant's Speech Processing

Given that infants demonstrate some sensitivity to potential prosodic markers of syntactic units, we can ask whether they use this information in organizing their representations of speech. With respect to the way prosodic bootstrapping unfolds, it could be that sensitivity to prosodic markers actually develops long before prosodic groupings play any role in organizing the information provided in fluent speech. Alternatively, it may be that infants are able to make immediate use of their sensitivity to prosodic markers as a means for organizing the input that they receive.

How can we determine when the organization that is potentially available in the prosody begins to play a significant role in speech processing? That is, how can we know whether or not infants are truly organizing the incoming speech signal into units such as clauses or phrases? The problem here is not far removed from one that early psycholinguistic researchers faced when trying to convince skeptical behaviorists that certain units of linguistic analysis corresponded to psychologically real processing units for listeners. The means by which the early psycholinguists were finally able to make their case was to show that linguistic units tend to be natural units for encoding and remembering information conveyed in speech.

A method used successfully with adults to demonstrate that linguistic units are actually used in on-line speech processing was to show that the organization provided by the linguistic structures had an impact on what information was remembered. Specifically, it was found that adults could better remember information from stimuli with a linguistic organization than with an arbitrary one (Marks and Miller 1964; Miller and Isard 1963; Suci 1967). Analogously, one can ask whether for the language learner, sensitivity to prosodic information also affords an organizational structure for encoding and remembering speech information.

Mandel, Jusczyk, and Kemler Nelson (1994) followed this research strategy in an investigation of whether infants are able to use prosodic organization to encode and remember speech information. Two-month-olds were tested using the version of the HAS procedure with the 2-minute delay between the preshift and postshift periods. To determine whether sentential prosody plays some role in organizing infants' memory for speech, Mandel and colleagues contrasted conditions in which such prosodic information was present to conditions in which it was not. In particular, they examined whether the phonetic properties of words that are prosodically linked within a single clause are better remembered by infants than the same words produced as individual items from a list. Mandel and coworkers reasoned that if prosody really helps in perceptual organization during on-line speech processing, memory for words should be better in the sentential context. Alternatively, there are certain perceptual grounds for actually predicting the opposite pattern of results (i.e., better performance in the list conditions than in the sentence conditions). This is because in some respects, the words in the list condition might be expected to include clearer phonetic cues. In particular, the words in the lists were produced one at a time and in citation form. By comparison, the phonetic characteristics of the individual words in sentential contexts are more apt to be influenced by the

phonetic characteristics of surrounding words. It is well known that words produced in fluent speech contexts are often less clearly articulated and more difficult to perceive when excised from context than the same words produced in citation form (Lieberman 1963; Pollack and Pickett 1964).

In their first experiment, Mandel and colleagues used three spoken sentences that had been recorded and selected from a larger group of unrelated sentences. The words in citation form were similarly selected from a longer spoken list of unrelated words. The words from the list were excised and rearranged to form three different word sequences that were otherwise identical to the three sentences. The overall durations of the list sequences were equated to the comparable sentences. Half of the infants in the study heard the sentences; the other half heard the lists. Thus, both groups of infants heard the same words in exactly the same order. The only difference was the prosodic envelope in which the words had been spoken.

During the preshift phase of the experiment, each criterion sucking response resulted in the presentation of either a single sentence or list sequence (e.g., "The rat chased white mice"). For a given infant, the same sentence or list sequence was played throughout the preshift phase. When the infant's sucking response habituated to this stimulus, the preshift phase ended and was followed by a 2-minute silent interval in which a series of colorful slides was presented. Then the postshift phase began. The infants heard either the same stimulus as in the preshift phase (control), one that differed by a single segment in one word (one phonetic change—e.g., "The cat chased white mice"), or one that differed by one segment in each of two words (two phonetic changes—e.g., "The cat raced white mice"). The results indicated that performance was significantly better for the sentential materials than for the lists. In particular, the infants responded with significant increases in sucking to both kinds of phonetic changes with the sentential materials, but not to either type of change with the list materials. Mandel and coworkers concluded that even 2-month-olds derive some benefit from the organization offered by sentential prosody in remembering speech information.

A follow-up experiment in the same study replicated and extended the findings by showing that 2-month-olds' memories for information within a sentence is better than for the same information in fragments of two adjoining sentences. Thus, infants were more likely to notice a change in "Cats like park benches" when this had been spoken as a single sentence than when the same words were excised from the sentences, "I know what cats like. Park benches are their favorite things to play on." Taken together, the results of the two experiments in this study suggest that the prosodic organization afforded by well-formed sentences facilitates infants' processing and memory for speech information.

More recently, the same investigators (Mandel, Kemler Nelson, and Jusczyk 1996) explored how prosodic structure may affect infants' abilities to better encode sequential order information. The ability to encode serial order is important for many different aspects of language acquisition. For example, serial order is important for the recognition of multisyllabic morphemes and words. The toddler learning English will find it valuable to distinguish between [ti' pat] and [pat' ti] ("teapot" and "potty"). An ability to encode serial order is also important for learning how word forms are formed and related to other words in the native language. For instance, derivational morphemes occur closer to word stems than do inflections (Kiparsky 1982; Mohanon 1986). In languages with highly developed morphological systems, such as Turkish, there are strict constraints on the ordering of sequences of morphemes within a word (Kenstowicz 1994). Finally, an ability to encode and remember serial order information may also be important for learning about syntactic relations within sentences and within phrases. For instance, changes in the ordering of words in English sentences often result in changes of meaning (compare "The dog bites the man" vs. "The man bites the dog"). In the long run, an inability to encode the order of lexical items would make it impossible for a learner to acquire a language like English, in which word order is used to encode many syntactic functions.

To determine whether prosodic organization helps infants in their encoding of the sequential ordering of speech sounds, Mandel, Kemler Nelson, and Jusczyk (1996) presented the same information either within a single well-formed prosodic unit or as fragments of two adjoining prosodic units. The sentence materials were created by having a naive female talker read the test sentences as part of a list of 15 unrelated sentences. The test sentences were "Cats would jump benches" and "Cats jump wood benches." The two sentences contained the same phonetic materials; only the ordering of the second and third words in each sentence changed. The creation of the comparable sentence fragment sequences was accomplished by having the same talker read a series of questions, followed by well-formed, but abbreviated two-word answers that might occur in the course of conversational speech. The answers used as stimuli were actually embedded in a larger set of questions and answers. The

question "Do cats jump or run?" was followed by "Cats jump," whereas the question, "Are those metal benches or wood benches?", was followed by "Wood benches." Subsequently, these sentence fragments were concatenated using a waveform editor to form the sequence "Cats jump. Wood benches." A different series of questions and two-word answers were used to prepare the sequence. "Cats would. Jump benches." Note that each two-word answer was a well-formed prosodic unit. However, when these two-word sequences were combined to form the four-word test sequences, they did not conform to sentential prosody, making them less prosodically coherent than the comparable sentential materials. Half of the infants were tested on the sentential materials and the other half were tested on the fragment sequences. During the preshift phase of the experiment, the 2-month-olds heard a stimulus such as "Cats would jump benches." After a 2-minute delay, the infants heard either the original stimulus or one in which a different ordering of the words occurred (e.g., "Cats jump wood benches"). Once again, infants performed significantly better when the information was presented within the same prosodic unit than when it occurred as parts of two different prosodic fragments. That is, only infants who heard the sentential materials reacted with significant increases in sucking to the word order changes. Therefore, these results provide another indication that the prosodic organization afforded in sentences does play some role in infants' processing and memory for speech information.

What these studies show, then, is that the prosodic packaging of clausal units seems to facilitate even very young infants' memory for speech information. What is not currently known is the extent to which the prosodic organization of units smaller than the clause might also influence language learners' encoding of and memory for speech. Given the likelihood that prosodic cues to subclausal units may be language-specific, and the data that suggest that infants do not show sensitivity to such units before 9 months of age, it does not seem promising to investigate whether 2-month-olds show similar memory benefits for information within prosodic phrases. Rather, it is important to develop procedures that could tap memory processes of older infants. One possibility is to adapt the methodology that Jusczyk and Aslin (1995) used to look at the detection of words in fluent speech. For instance, if infants were familiarized with information that occurred within a particular clause or phrase as opposed to information that occurred in fragments of clauses or fragments of phrases, would they be more apt to recognize this material when it appeared later in a longer passage? Kemler Nelson, Mandel, and I have tried this for clausal units with 6-month-olds and find that they do perform better when they have been familiarized with material that occurred as a well-formed clausal unit, as opposed to fragments of two different units. Thus, we were able to replicate the findings that we had obtained with 2-month-olds in this older age group and with a new procedure. The next step is to test 9-month-olds with material within phrasal units, which we hope to do in the future.

Reassessing Prosodic Bootstrapping

Considering the evidence reviewed in the preceding three sections, how far can prosodic bootstrapping take the language learner? What the findings show is that clausal units are well marked in the prosody of utterances, that infants are sensitive to this marking, and that it appears to play some role in their encoding and retention of speech information. There is some evidence as well that subclausal units such as major phrases receive prosodic marking and that infants are sensitive to this when it occurs in the input. However, the marking appears to be primarily directed toward prosodic phrase boundaries rather than syntactic ones per se (see Gerken 1996 for a similar view). Moreover, when prosodic and syntactic phrase boundaries do not correspond to each other, it appears to be the prosodic phrase boundaries that young language learners respond to.

One interpretation of this pattern of findings is that prosodic bootstrapping accounts help to explain how infants correctly locate clausal units in the input, but nothing beyond that. This view is based largely on the assumption that if the marking of syntactic phrase boundaries by prosodic cues is less than perfect, then prosodic bootstrapping accounts must fail. Moreover, the fact that on any given occasion, a putative prosodic marker of syntactic boundaries (e.g., syllable lengthening) can also be used for some other communicative function or purpose is seen as problematic. How is the infant to know whether, on a given occasion, the prosodic cue is marking a syntactic boundary or something else?

The position just outlined is probably overly pessimistic for a number of reasons. Consider the last sort of objection, that any particular cue has various uses on different occasions. In fact, the different prosodic indices of phrasal boundaries often work in combination with one another rather than singly. Because these cues are multiple and tend to covary, ambiguity is less troublesome than it would be if the cues were single. Moreover, prosodic bootstrapping is not equivalent to reading the syntactic

organization directly from the prosody. Rather, prosodic bootstrapping provides the learner with a kind of rough grouping of elements in the input that then positions the learner to pull out the underlying syntactic organization. Hence, it may be equivalent to the kind of rough categorization of speech sounds that very young infants have before their perceptual categories are modified by language-specific input. Consequently, the prosodic groups do not have to correspond to the syntactic groups on every occasion to be helpful in discovering the underlying syntactic organization of native-language utterances. Just the mere grouping of the input into smaller processing units may prove helpful in itself. The process may facilitate further analyses in much the same way as the 7,5-month-old English-learner's use of a metrical segmentation strategy for word segmentation does (i.e., as when the latter strategy provides smaller units of analysis that then allow for a distributional analysis of phonotactic and allophonic cues to word boundaries). The analogous situation in the present case is that prosodic phrase groupings, combined with infants' developing word segmentation abilities, provide the means to work out the distributional patterns of words within such prosodic units.

Finally, as Fisher (in press-a) has noted, language learners typically receive sentence fragments in addition to whole sentences in the input. Most of these are well-formed phrasal units, which the learner could use in cross-sentential comparisons to derive information about syntactic constituents in utterances. For example, the child who hears "the chocolate cake" in response to a question may subsequently be able to use this information in figuring out the constituents of the sentence "She dropped the chocolate cake." Furthermore, it is worth noting that the kinds of fragments that are *not* well-formed phrasal units (viz., errors, hesitations, interruptions, etc.) are precisely those kinds of utterances that are likely to be marked with prosody that suggests that these are not complete units. The prosodic features of such utterances are likely to be similar to what occurs in the noncoincident samples in the studies of perception of phrasal units by infants, namely, they have pauses in places not predicted by the other prosodic correlates to phrasal boundaries.

Before turning to a discussion of other kinds of information in the speech signal that may prove useful for learning about syntactic organization, there is another point that should be considered. Specifically, what is the relationship between sensitivity to prosodic groupings and the capacity for segmenting words from utterances? At one point, Kemler Nelson, Hirsh-Pasek, and I (Kemler Nelson et al. 1989) suggested that sensitivity to units in the input might follow some sort of differentiation process from larger to smaller. That is, sensitivity would develop first for clausal units, then for phrasal units, and ultimately for word units in the input. Indeed, such a developmental sequence is suggested by the pattern of results in our studies (e.g., Hirsh-Pasek et al. 1987; Jusczyk et al. 1992a; Myers et al. 1996) that measure preferences for speech in which pauses have been inserted at or within unit boundaries.

However, the development of new test procedures has produced new evidence that has caused us to reconsider our account. In particular, there is now ample reason to believe that infants begin some sort of rough word segmentation process before they exhibit sensitivity to prosodic markers of phrasal groupings in utterances. Thus, what seems to be happening is that two sorts of processes are going on at the same time with respect to how the infant analyzes the speech signal. On the one hand, the infant is discovering important ways information is grouped or clustered in utterances (i.e., into clauses, and subsequently, into phrases). On the other hand, the infant is also discovering the elementary units (i.e., words) that function inside such larger groupings of information. This latter process that may yield some wordlike units (based on metrical stress properties) at 7,5-months and more precise word segmentation (perhaps based on pho-notactic and allophonic properties, among other things) over the course of the next three months or so. Ultimately, both kinds of processes can have some bearing on the discovery of the syntactic organization of the native language. To this point, we have focused on the first of these processes— prosodic cues to the way that information is grouped within utterances. Let us now explore how other information in the speech signal may provide clues to syntactic organization.

Other Information in the Signal that May Facilitate the Acquisition of Syntax

We have been considering the way prosody could help to mark breaks in utterances that could reflect their underlying syntactic organization. However, information about specific sound patterns and where these occur could also provide hints about the nature of syntactic constituents. For example, correlations between certain types of sound patterns of words and their grammatical categories could be potentially helpful in acquiring syntax. Kelly (1992) has noted a number of ways information in the speech signal may be correlated with grammatical category assignment (see also Sereno and Jongman

1995). For example, in bisyllabic English nouns, stress typically occurs on the first syllable, whereas in bisyllabic English verbs, stress is often on the second syllable. In fact, in a survey of over 3,000 nouns and 1,000 verbs in English, Kelly and Bock (1988) found that 94 percent of the nouns had first-syllable stress and 69 percent of the verbs had second-syllable stress.

These distributional properties have some real consequence for fluent English-speakers' behavior: Subjects were more likely to pronounce pseudowords with stress on the first syllable if they thought that these were nouns and on the second syllable if they thought these were verbs. In addition, Cassidy and Kelly (1991) have observed evidence for a significant relationship between syllable number and grammatical class in parental speech to 15-month-old infants. Nouns generally had more syllables than did verbs. In fact, the likelihood that a given word was a noun was 38 percent for one-syllable words, 76 percent for two-syllable words, 92 percent for three-syllable words, and 100 percent for four-syllable words in the input corpora that they analyzed. Obviously, in order to use these kinds of phonological properties to differentiate among nouns and verbs in the input, the learner must have already stored some exemplars of each type of category, along with the frequency with which certain sound patterns are associated with them. One way this may happen is that when words begin to be acquired, learners track the sound properties of words that refer to objects separately from those that refer to actions. This strategy would allow them to note any differences in the frequency with which certain sound patterns are associated with these categories. As the investigations of the perception of phonotactic patterns demonstrate (e.g., Jusczyk et al. 1994), infants are very sensitive to the frequency with which certain patterns appear in the input. Once these kinds of relationships are detected with respect to the first few nouns and verbs in the lexicon, they could be applied to routines that are used to analyze new input strings.

In addition to potential differences in the sound patterns of nouns and verbs, many investigators have focused on similar sorts of distinctions among content and function words. For example, Jakobson and Waugh (1987) observed that function morphemes from the same language usually share certain phonological properties. Function words are also usually shorter than content words, and they often have different phonemes, such as [ð] in English (Gerken 1996). Moreover, content words are more likely to be stressed than are function words (Gleitman et al. 1988; Kelly 1992). Thus, infants could potentially use such differences as a means of distinguishing function and content words. Morgan and his colleagues have explored a whole range of acoustic properties that could serve to differentiate function words from content words across a range of different languages (Morgan et al. 1996). They did not find evidence that any one property was strongly correlated with either class of words. Nevertheless, they found that function and content words did differ significantly in their distributions of such cues, such that the combination of several such cues was predictive of which class a particular item was likely to belong to. Moreover, in analyzing input to English-learning children, they found that the words with the highest frequencies of occurrence tended to be function words.

In considering these findings on the relation between phonology and grammatical categories, it is important to keep in mind that what the sound properties can provide is a rough categorization of the input. Counterexamples will inevitably arise. Sound properties are surely not perfect predictors of a word's grammatical category. But once again, they may help in a first-pass analysis; they may enable the learner to correctly categorize the information on a number of occasions and then to use other sources of information (semantic, syntactic, pragmatic) to eventually correct those erroneous categorizations that do occur. Moreover, some of these other sources of information may, in the long run, turn out to be more reliable than the information regarding the sound properties of words. If so, we should expect to see a greater reliance on these other lands of properties as the infant's grasp of the different levels of organization of the native language increases during the course of development. Thus, the use of these kinds of sound relations may be an interim strategy until the infant's skills are better developed and more established. This, after all, is what bootstrapping is all about. It is a means to get the system started in the right direction. The routes that were used early on need not be involved in how the system functions in adults (Bever.1975) other than to show up occasionally in some psycholinguistic task, such as judging whether a particular pseudoword is more likely to be a noun or a verb in the absence of any semantic information. This may be another example of Newport's "less is more" principle at work. As processing resources come on line and new skills are developed, the infant has more possibilities to choose from. The kinds of information that were relied on heavily when resources were scarce are used less often as other alternatives increase. Thus, infants could at times be more sensitive to the presence of certain kinds of acoustic information than are adults because the

greater processing resources of the latter group have led them to rely on other kinds of information.

We have been considering the way the presence of certain kinds of sound patterns could play a role in facilitating the assignment of grammatical categories to words. However, there are other ways simply learning to recognize the sound pattern of certain words could help in discovering the syntactic organization of utterances. By identifying and storing information about certain sound patterns and by being able to recognize these when they occur in fluent speech, the infant is in a position to observe how these are distributed within the groupings that they extract from utterances. In English, frequently occurring articles such as "the" or "a" rarely occur at the ends of such units, but they commonly occur at the beginnings of these. Moreover, they occur prior to nouns and adjectives, but not immediately before verbs. Thus, by restricting distributional analyses to co-occurrences involving the phonologically defined set of grammatical morphemes and the content words with which they co-occur, learners could potentially distinguish among syntactic phrase types.

One possible problem with the picture that was just presented is the widespread belief that language learners have difficulty perceiving function words because they are unstressed (Echols and Newport 1992; Gleitman et al. 1988). If such words are difficult for language learners to detect in fluent speech, then it is hard to see how they could play much of a role in facilitating a distributional analysis of the input. However, several recent investigations suggest that infants are sensitive to the occurrence of function words in fluent speech. Initial support for such a view comes from a study demonstrating that by 11 months, infants appear to be sensitive to the phonological contribution that grammatical morphemes make to the overall "sound" of their language (Shafer et al. 1992). In particular, Shafer and coworkers found that infants could distinguish a normal English passage from one in which nonsense syllables that were phonologically unlike English grammatical morphemes replaced a subset of actual morphemes.

In another study, Shady, Gerken, and Jusczyk (1995) investigated whether 10,5-month-olds show any recognition of the correct ordering of function and content words in their native language. They modified a series of speech passages by interchanging the order of one function-content word pair in each sentence (e.g., "She her felt mother's soft, warm fur," where "her" and "felt" were reversed). The original and modified passages were synthesized using a DECTalk speech synthesizer to ensure that the prosody was not unnatural in the case of the modified samples. When the infants were tested on both versions of these synthesized passages with the headturn preference procedure, they listened significantly longer to the original than to the modified passages. Inspection of the individual passages that were used in testing indicated several kinds of cues that infants may have used to detect the word order reversals. For example, in a number of the modified passages, reversing the order of function and content words led to a sequence of two function morphemes in a row (a sequence that is relatively infrequent in English sentences). Further research is necessary to determine precisely just which sources of information Shady and colleagues' infants were responding to. In any case, English-learning l0,5-month-olds seem to have discovered some source of information about how function and content words are typically ordered in utterances in their native language.

The findings reviewed in this section point to other ways information that is available in the acoustic signal could be involved in discovering the syntactic organization of one's native language. The sensitivity that infants display to the recurrence of certain sound patterns in the input, plus their apparent ability to detect the occurrence of words in fluent-speech contexts, are two factors that may position them for a distributional analysis of the input. What is needed at this point is more detailed information about when particular words (especially function words) are recognized by infants. Also, it would be interesting to know whether infants display sensitivity to the kinds of differences that are potentially correlated with grammatical classes of words. A better understanding of these matters would help delineate how much the acquisition of syntax is facilitated by information in the speech signal. Finally, since our focus is on acquisition of the sound structure of the language, we have concentrated on the kind of bootstrapping that could be derived from information in the signal. This does not preclude the possibility that bootstrapping of other sorts also occurs during language acquisition. In fact, it seems very likely that the language learner draws on information from many other sources in order to determine the organization of utterances in the native language.

Exploring Relations between Sounds and Meanings

At some point, in order to fully understand and explain language acquisition, it is necessary to

determine how processes at various levels of linguistic organization interact during the course of development. For example, although it is reasonable to suppose that the most important factors influencing whether or not a child will learn a new word meaning have to do with semantic organization (and the child's current cognitive state), other levels of language organization could also affect the process. Thus, the achievement of an optimal way of processing information at some level of linguistic organization (e.g., phonological, syntactic, semantic) may well constrain the possible forms of organization at other levels. That is, the language learner has to find the path that simultaneously satisfies best the task demands at all levels of linguistic organization.

One area in which influences from different levels of organization might be expected is in the development of a lexicon for the native language. Current views are that lexical entries contain information from a number of different levels of organization (e.g., see the papers in Gleitman and Landau 1994). In addition to some description of the sound properties of a lexical item and its associated meanings, an entry may include information about the word's syntactic category and the kinds of syntactic structures that it can participate in. The assumption that lexical entries are potentially rich in information raises some interesting questions about how the lexicons in general and lexical entries in particular develop. For example, do lexical entries all begin in the same way, perhaps as meanings that then get associated to sound patterns, with syntactic information only added at some subsequent point? Are the only limitations on the addition of new lexical entries ones that have to do with whether or not some particular meaning already has a sound label attached to it—that is, something along the lines of Markman's (1989; 1991) mutual exclusivity principle? Does the nature of the sound pattern of a particular word ever affect the likelihood of its being added to the lexicon?

Some Influences of Naming on Categorization Behavior

For the moment, let us consider how the sound structure of words could potentially impact on lexical growth. One possibility is that the presence of a verbal label influences the learner's categorization of objects and events. By now there are many demonstrations that the nature of verbal labels influences the categorization behavior of older infants. For example, 2- to 3-year-olds have been shown to interpret novel count nouns as referring to basic level and superordinate level categories of objects (D'Entremont and Dunham 1992; Markman and Hutchinson 1984; Soja, Carey, and Spelke 1991; Waxman and Kosowski 1990). However, when novel adjectives are used, children appear to interpret them as referring to object properties or to subordinate level distinctions (e.g., Gelman and Markman 1985; Hall, Waxman, and Hurwitz 1993; Katz, Baker, and Macnamara 1974; Smith, Jones, and Landau 1992; Taylor and Gelman 1988). Moreover, just the mere fact that a name is used in a situation has been shown to influence children's categorizations of objects (Gelman and Taylor 1984; Golinkoff et al. 1992; Markman and Hutchinson 1984; Waxman and Gelman 1986). For example, in their investigation, Markman and Hutchinson (1984) found that 2- and 3-year-olds who heard an object labeled with a term like "dax" were much more likely to respond to a request to "Find another dax" by choosing an object with similar properties (i.e., a taxonomic choice) than one that bore a thematic relation to the first named object (i.e., a thematic choice). By comparison, when the label was not used in the request (i.e., "Find another one"), the same children were more likely to make a thematic choice. More recently, Landau and Shipley (1995) reported results of an investigation that demonstrated that the number of different labels used (one vs. two) influences whether children as young as 2 years are likely to group objects into a single category or not.

Furthermore, the impact that naming an object has on categorization performance has been reported for 1-year-olds and even younger infants. In a series of studies, Waxman and Markow (1995) found that 12-month-olds responded differently on a categorization task depending on the availability of a verbal label. Half of the infants heard an experimenter label an object using a novel noun phrase (e.g., "Look, an animal"); the other half a heard a more general phrase (e.g., "Look, what's here"). During a subject test period, the infants who heard the noun phrase were significantly more likely to distinguish between a new instance from the familiar category and an instance from a novel category. Waxman and Markow interpreted this finding as an indication that the availability of the same label "invites infants to search for coherence among the different objects and to form object categories."

Balaban and Waxman (1995) extended these findings in several interesting directions. First, they demonstrated that infants as young as 9 months of age are affected by the availability of a verbal label on a categorization task. Second, they demonstrated that the changes in categorization behavior by the presence of a verbal label involve more than a general alerting effect provided by auditory stimulation.

In particular, they found differential patterns of responding when a verbal label (even a low-pass-filtered one) was used during a familiarization period with an object as compared to when a nonspeech auditory stimulus (a sinewave tone matched to the label in loudness and duration) was used. Novelty preferences during the test period were greater when infants had heard the words rather than the tones during familiarization. Thus, there are some indications that the presence of verbal labels are beginning to influence categorization behavior in language learners around the time when they begin to show the first signs of comprehension of words (e.g., Benedict 1979; Huttenlocher 1974). Infants at this age appear to be primed to begin to link verbal labels to their categorizations of objects and events in their immediate environment.

Sound Patterns and Representations in the Lexicon

Investigations by child phonologists (e.g., Ferguson and Farwell 1975; Schwartz 1988; Vihman et al. 1985) suggest another way the sound structure of words can affect the pattern of lexical growth—namely, infants may avoid words that contain sounds that are difficult for them to produce. However, as Schwartz has noted, although this tendency would be expected to affect the nature of the lexicon underlying the child's production of words, there is no necessity that it should affect the words in the child's receptive lexicon. Another possibility is that the infant might at least initially avoid adding to the lexicon words that are difficult to discriminate from existing items in the lexicon. For example, models such as WRAPSA (Jusczyk 1993a) and CHIPHO (Suomi 1993) assume the representations of sound patterns of early lexical items are not fully detailed descriptions. Consequently, if many words with very similar sound patterns are added at an early stage, then the lack of sufficient detail in the representations might lead the infant to recognition errors. In fact, studies based on estimates of the vocabularies of children 5 to 7 years old suggest that children's lexicons have many fewer lexical neighbors than the same words in adults' lexicons (Charles-Luce and Luce 1990, 1995; Walley 1993; but see Dollaghan 1994). This finding is consistent with the view 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.