Infants' Preference for the Predominant Stress Patterns of English Words

Peter W. Jusczyk

State University of New York at Buffalo

Anne Cutler

MRC Applied Psychology Unit, Cambridge, U.K.

Nancy J. Redanz

State University of New York at Buffalo

Jusczyk, Peter W.; Cutler, Anne; and Redanz, Nancy J. *Infants' Preference for the Predominant Stress Patterns of English Words*. Child Development, 1993, 64, 675–687. One critical aspect of language acquisition is the development of a lexicon that associates sounds and meanings; but developing a lexicon first requires that the infant segment utterances into individual words. How might the infant begin this process? The present study was designed to examine the potential role that sensitivity to predominant stress patterns of words might play in lexical development. In English, by far the majority of words have stressed (strong) initial syllables. Experiment 1 of our study demonstrated that by 9 months of age American infants listen significantly longer to words with strong/weak stress patterns than to words with weak/strong stress patterns. However, Experiment 2 showed that no significant preferences for the predominant stress pattern appear with 6-month-old infants, which suggests that the preference develops as a result of increasing familiarity with the prosodic features of the native language. In a third experiment, 9-month-olds showed a preference for strong/weak patterns even when the speech input was low-pass filtered, which suggests that their preference is specifically for the prosodic structure of the words. Together the results suggest that attention to predominant stress patterns in the native language may form an important part of the infant's process of developing a lexicon.

The process of understanding speech involves recognizing the individual words of which an utterance is composed. Language users cannot store in memory every complete utterance that might be presented to them. George Miller (1964) spelled out the impracticality of any scheme of this sort. Aside from clichés, most utterances represent novel combinations of words. The number of permissible English sentences of 20 words or less is on the order of 10^{20}. It would take about 100,000,000,000 centuries simply to utter these, let alone learn them by rote. What is stored, therefore, cannot be whole utterances; instead, it must be the discrete units—words—of which utterances are composed. For mature language users, the processing of speech input thus involves recognizing in the input the sound patterns that correspond to these discrete lexical units.

But how does this process begin? How does an immature language user start to perceive words? The infant may well come into the world armed with an expectation that there will be words, that is, that speech patterns will map onto the world in discrete chunks. Certainly, there is abundant evidence that human speech is attractive to infants from a very early age (Colombo & Bundy, 1981; Friedlander & Wisdom, 1971; Glenn, Cunningham, & Joyce, 1981), and

especially child-directed speech (Cooper & Aslin, 1990; Fernald, 1985; Fernald & Kuhl, 1987; Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978; Pegg, Werker, & McLeod, in press). Nevertheless, it cannot be the case that infants come into the world with a preprogrammed expectation of what words will be like, because the structure of words differs across languages, yet any infant can acquire any human language to which he or she is exposed. Thus finding out what words are like must form part of the very task of language acquisition.

As Mehler, Dupoux, and Segui (1990) have noted, finding words in the speech stream is not a simple task for the child, because the input does not consist solely of isolated words, even in child-directed speech. In any case, child-directed speech is only a subset of what the infant hears in his or her environment. The problems of identifying boundaries between words in continuous speech are well known to speech researchers. Klatt (1989) has described many of the potential pitfalls posed by the fact that in continuous speech there are both multiple alternative starting points for words, and also multiple alternative pronunciations of words created by phonological processes operating across word boundaries. These kinds of problems make it extremely unlikely that simply relying on information about phoneme sequences would enable a listener to segment continuous speech into words. The problems are, of course, further compounded by the fact that speech is frequently rendered less than fully clear, by background noise or other factors. Harrington and Johnstone (1987) have computed the possible divisions into words of sentences presented in incompletely specified phonetic transcription; they found that even sentences of about seven words could often represent millions of alternative possible word strings.

Given that the problem of segmenting continuous speech into words is nontrivial even for mature listeners, how does the infant begin to solve the problem? The infant has to find out how to divide the continuous speech stream into the lexical units that the particular language consists of, without any knowledge at all of what these lexical units are like. One proposal is that the infant can acquire cues to lexical segmentation by paying attention to the prosodic characteristics of the input language (Gleitman, Gleitman, Landau, & Wanner, 1988; Hirsh-Pasek

et al., 1987; Jusczyk, 1991, 1992). Certainly it is true that the prosodic structure of language is highly salient to the infant from a very early age—for instance, newborn infants show sensitivity to gross features of the prosodic structure of the native language (Mehler et al., 1988). They can even discriminate at this early age certain prosodic correlates (in French words) of the presence versus absence of a word boundary (Christophe, Dupoux, Bertoncini, & Mehler, 1993). Similarly, recent evidence suggests that by 4½ months American infants show sensitivity to clause boundaries in English (Jusczyk, 1989), and that by 9 months, they are showing sensitivity to boundaries of major phrases (Jusczyk et al., 1992). In both cases, there are indications that infants are responding to prosodic features in the input, because the sensitivity is evident when the speech input is low-pass filtered to remove segmental information (while preserving prosodic information). Of course, in addition to this work, there is a growing body of evidence suggesting that infants are learning to differentiate native from foreign language input during the latter half of the first year of life (e.g., Best, 1991; Best et al., 1990; Werker & Lalonde, 1988; Werker & Tees, 1984). Moreover, there are also indications that infants are beginning to pick up information about properties particular to their own native language during this same period. For example, Jusczyk, Friederici, Wessels, Svenkerud, and Jusczyk (1993) recently demonstrated that 9-month-olds, but not 6-month-olds, appear to be sensitive to phonotactic constraints on words in the native, as opposed to a foreign, language. Phonotactic constraints refer to the restrictions that exist on the way that phonetic segments can be ordered within words in a language. Sensitivity to another property of the native language input, the structure of its vowel categories, is apparently present in infants as young as 6 months of age (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). Thus during the latter half of the first year, infants are beginning to absorb information about distributional properties of the native language input.

We expect that one of the aspects of words to which infants should pay attention is their characteristic prosodic structure. In English, the major prosodic feature applying at the word level is word stress. As in other stress languages, there are in English two types of syllable in polysyllabic words: strong syllables (bearing primary or second-

ary stress and containing full vowels), and weak syllables (unstressed, and containing reduced vowels). The characteristic stress rhythm of English is an alternation of strong and weak syllables, and this prosodic structure could be highly salient even to infants. Certainly there is evidence that infants can make discriminations based on stress (Jusczyk & Thompson, 1978; Karzon, 1985; Spring & Dale, 1977). For instance, Jusczyk and Thompson tested whether 2-month-old infants could distinguish utterances that contrasted only in their stress patterns. They found that infants easily discriminated utterances like [ba' da] from [ba da']. Hence it is clear that infants at least have the ability to discriminate the acoustic correlates of word stress differences from an early age.

The possible arrangements of strong and weak syllables within words are not equally represented within the lexical repertoire of English. The most common word type in English is a bisyllable with a strong initial syllable and a weak second syllable (Carlson, Elenius, Granstrom, & Hunnicutt, 1985). Only about a quarter of the words of English are weak-initial polysyllables (Cutler & Carter, 1987). Moreover, because many weak-initial words have a low frequency of occurrence, they are underrepresented in actual speech in comparison to their representation within the vocabulary; in a corpus of 200,000 words of spontaneous British English conversation, Cutler and Carter (1987) found that weak-initial polysyllabic lexical (or "content") words accounted for only about 4% of all words. Thus there is a marked asymmetry, both in the vocabulary and even more so in natural speech, with regard to how often particular stress patterns occur. Strong-initial words (monosyllables and initially stressed polysyllables) are likely to be heard often; weak-initial words are likely to be heard only rarely. In fact, weak-initial sequences in by far the majority of cases consist of an unstressed grammatical word plus a following lexical word—that is, such sequences need to be segmented for lexical access.

Speech to children exaggerates this feature of English as it exaggerates other prosodic characteristics of the language (Fernald & Simon, 1984). Kelly and Martin (in press) report that the relative frequency of strong onsets in a sample of speech to children was even greater than the relative frequency of strong onsets in Cutler and Carter's (1987) corpus study.

The possibility exists, therefore, that infants may be able to exploit prosodic information efficiently in solving the segmentation problem in English. There is evidence that infants can discriminate stress patterns; and distributional studies show that stress patterns are very asymmetrically distributed in the English vocabulary. Do infants perceive that this asymmetry exists in the repertoire of English words, and do they use this knowledge in building their vocabulary? The present studies were designed as an initial approach to answering this question.

Experiment 1

In previous investigations (e.g., 1992, 1993), we have found indications that infants at around 9 months of age demonstrate sensitivity to structural features of the sound patterns found in their native language. Accordingly, we began our investigation of infants' sensitivity to the predominant stress patterns of English words with 9-month-olds. Our aim was to determine whether American 9-month-olds display a preference for listening to words that follow the predominant strong/weak pattern as opposed to the less frequent, but allowable, weak/strong pattern. One possible indication of such a preference would be if infants orient significantly longer to words following a strong/weak pattern than they do to ones following a weak/strong pattern. To explore this possibility we used a headturn preference paradigm, first developed by Fernald (1985) and subsequently modified and used in studies that we have conducted on infants' sensitivity to perceptual units in their native language (e.g., Hirsh-Pasek et al., 1987; Jusczyk et al., 1992).

Method

Subjects.—Twenty-four infants of approximately 9 months of age (12 males and 12 females) were tested from the suburban Buffalo area. The infants had an average age of 39 weeks, 0 days (range: 34 weeks, 4 days to 41 weeks, 6 days). Eight additional infants were tested but were not included for the following reasons: failed to look for an average of at least 3 sec to each side (3), parent failed to center the infant on his or her lap (4), English was not the primary language spoken at home (1).

Stimulus materials.—The materials consisted of 16 prerecorded lists of English words. Each list consisted entirely of 12 two-syllable words. In half of the lists, the words all followed a strong/weak syllable accent

pattern. In the other half of the lists, the words all followed a weak/strong accent pattern. An example of each type of list is shown in Table 1 (the complete set of materials is listed in the Appendix). In generating the lists, the strong/weak and weak/strong words were matched in terms of the vowel that occurred in the stressed syllable. In addition, an effort was made to match as much of the phonetic material in the rest of the words as closely as possible. For example, a word with a strong/weak pattern, like "crossing," was matched to one with a weak/strong pattern, like "across." Once the 96 words of each stress type were selected, they were divided into eight lists of 12 items each.

The lists were recorded in a sound attenuated room with a Shure microphone (SM10A) and a Revox (A77) tape recorder. The talker was a college-aged, female, native English speaker from western New York. She practiced reading the lists several times before beginning the recordings. In recording the lists, the talker alternated between lists of strong/weak and weak/strong items. The talker was told to read each list at a comfortable rate and to try to space the words equally. One version of each of the prerecorded lists was then digitized into a separate file on a VAXStation 3176 using a 12-bit A/D converter. Measurements of the durations of each list were made. The weak/strong lists had an average duration of 14.85 sec and the strong/weak lists had an average duration of 14.65 sec. A t test indicated that the two types of lists did not differ significantly in their overall durations, $t(7) = 0.694$. Two lists from each stress pattern type were chosen as practice trial lists; the remaining lists were used on the test trials.

TABLE 1

Sample Lists

W/S	S/W
comply	pliant
befall	falter
condone	donor
comport	comet
pomade	neighbor
abut	butter
define	final
restore	stalwart
resent	gentle
assign	sinus
caprice	rhesus

Acoustic measurements (pitch, amplitude, and duration) were taken of all 144 test words. As expected, there were significant differences between the stressed and unstressed syllables. Stressed syllables were longer (mean duration = 326 ms) than unstressed ones (mean duration = 182 ms), stressed syllables showed greater pitch range (mean = 52 Hz) than unstressed ones (mean = 46 Hz), and the mean amplitude of stressed syllables was greater (76 rms amplitude) than that of the unstressed ones (68 rms amplitude). The strong/weak and weak/strong word sets did not differ significantly on any amplitude measure, but they did differ significantly in duration and pitch range. The weak/strong words (mean duration = 546 ms) were significantly longer than the strong/weak words (mean duration = 471 ms; $t[142] = 6.55$, $p < .001$), and the pitch range of the stressed syllables was significantly greater for the weak/strong words (60 Hz) than for the strong/weak words (44 Hz; $t[142] = 2.11$, $p < .04$). Prosodic effects are therefore more marked in the weak/strong word set than in the strong/weak set. Note that child-directed speech exhibits exaggerated prosodic effects (Fernald & Simon, 1984), so that, if anything, our weak/strong word set resembles child-directed speech to a greater degree than does our strong/weak word set.

Apparatus.—The digitized files were transferred from the VAXStation to a PDP 11/73 computer. During the experiment, the PDP controlled the presentation of the lists and recorded the observer's coding of the infants' responses. The audio output for the experiment was generated from the digitized waveforms of the samples. A 12-bit D/A converter fed the output through anti-aliasing filters and a Kenwood audio amplifier (KA 5700) to 7-inch Advent loudspeakers mounted on the side walls of the testing booth.

The experiment was conducted in a three-sided test booth constructed out of pegboard, with panels of 4 × 6 feet on three sides and open at the back. This made it possible for an observer to look through one of the existing holes to monitor the infant's headturns. Except for a small section for viewing the infant, the remainder of the pegboard was backed with white cardboard to guard against the possibility that the infant might respond to movements behind the panel. The test booth had a red light and a loudspeaker mounted at eye level on each

of the side panels, and a green light mounted on the center panel. A white curtain suspended around the top of the booth shielded the infant's view of the rest of the room. A computer terminal and response box were located behind the center panel, out of view of the infant. The response box, which was connected to the computer, was equipped with a series of buttons that started and stopped the flashing center and side lights, recorded the direction and duration of head-turns, and terminated a trial when the infant looked away for more than 2 sec. Information about the direction and duration of headturns and the total trial duration was stored in a data file on the computer.

Procedure.—The procedure was a modified version of one originally developed by Fernald (1985). Each infant was held on a parent's lap. The parent was seated in a chair in the center of the test booth. The infant completed a four-trial familiarization phase (two lists of each type: strong/weak and weak/strong) and a 12-trial test phase. The weak/strong lists were consistently played through the loudspeaker on one side panel and the strong/weak lists through the loudspeaker on the other side panel. (The side was counterbalanced across subjects.) The familiarization phase was intended to acquaint the infant with the assigned position of each type of list. The ordering of the stimulus lists during the test trials was random, subject to the constraint that no more than three lists of the same type could occur in a row. Each subject was tested with a different random ordering of the lists.

Each trial began by blinking the green light on the center panel until the infant had oriented in that direction. Then, the center light was extinguished and the red light above the loudspeaker on one of the side panels began to flash. When the infant made a headturn of at least 30° in the direction of the loudspeaker, the next list appropriate to that side began to play and continued until its completion or until the infant failed to maintain the 30° headturn for 2 consecutive sec (e.g., if the infant turned back to the center or the other side, looked at the mother, the floor, or the ceiling). If the infant turned briefly away from the target by 30° in any

direction, but for less than 2 sec, and then looked back again, the time spent looking away was not included in the orientation time. During the familiarization trials, the red light was extinguished when the list began, but during the test trials the light remained on for the entire duration of the trial.[1]

An observer hidden behind the center panel looked through a peephole and recorded the direction and duration of the infant's headturns using a response box. The observer was not informed as to which loudspeakers played the strong/weak and weak/strong lists. This was possible because the assignment of the versions to the left or right side was determined by the computer and not revealed to the observer until the completion of the test session. The loudness levels for the samples were set by a second assistant, who was not involved in the observations, at 72 ± 2 dB (C) SPL using a Quest (Model 215) sound level meter. In addition, both the observer and the infant's parent listened to a recording of lists of randomly interspersed strong/weak and weak/strong words over headphones. This list proved to be an excellent masking stimulus, and parents and observers reported that with this background they were unaware of the nature of the stimulus at a given location on a particular trial.

Results and Discussion

The amount of time that each infant oriented to the loudspeaker on each trial was recorded. The average looking times were 5.43 sec (SD = 1.48 sec) for the weak/strong lists and 7.45 sec (SD = 2.09 sec) for the strong/weak lists. Twenty-one of the 24 infants had longer looking times for the strong/weak lists. A paired t test confirmed that the difference in orientation times to the strong/weak and weak/strong lists was significant, $t(23) = 4.39$, $p < .001$.

Thus, the results indicate that 9-month-old American infants do show a preference for the lists of words following a strong/weak stress pattern. This suggests that they may already have developed some sensitivity to the dominant stress pattern of English words. Of course, one could raise the ques-

[1] Extensive pilot testing before previous studies (e.g., Jusczyk et al., 1992) convinced us that this is the best way to handle the lights during the procedure. Leaving the flashing light on during the familiarization trials seems to habituate the infants to the lights, and results in very short orientation times during the test trials. Moreover, the infants are also less likely to complete the full set of test trials under these circumstances.

tion as to whether the preference for strong/ weak patterns reflects attention to the most frequently used word stress pattern in the language or whether words with these patterns are simply more intrinsically interesting or pleasing to the infants. For instance, it may be the case that infants of any age might show the same preference for the strong/weak patterns. If so, then the present results would have little to say about infants' learning about the sound patterns of their native language. To explore this possibility, we decided to test a group of younger infants on the same materials.

Experiment 2

In previous investigations of infants' sensitivity to structural features of the sound patterns of their native language, there were indications that although 9-month-olds displayed such sensitivities, 6-month-olds often did not (e.g., Jusczyk et al., 1992, 1993). Moreover, the headturn preference procedure has also been used successfully to demonstrate that 6-month-olds show sensitivity to prosodic correlates of syntactic units in the native language (Hirsh-Pasek et al., 1987) and to musical phrase structure (Krumhansl & Jusczyk, 1990). For these reasons, we next tested a group of 6-month-olds on the strong/weak and weak/strong lists. An indication that this younger group also demonstrates a preference for the strong/weak lists would support the view that this type of pattern is simply inherently more interesting for the infants.

Method

Subjects.—Twenty-four infants of approximately 6 months of age (15 males and nine females) from the suburban Buffalo area were tested. The infants had an average age of 27 weeks, 2 days (range: 22 weeks, 0 days to 29 weeks, 3 days). An additional three infants were tested but not included for the following reasons: crying (1), experimenter error (1), and difficulty turning to one of the two sides (1).

Stimulus materials, apparatus, and procedure.—These were the same as in the previous experiment.

Results and Discussion

The amount of time that each infant oriented to the loudspeaker on each trial was recorded. The average looking times were 7.68 sec (SD = 2.11 sec) for the weak/strong lists and 7.69 sec (SD = 1.66 sec) for the strong/weak lists. Thirteen of the 24 infants had longer average looking times for the

strong-weak lists. A paired t test indicated that the difference in orientation times to the strong/weak and weak/strong lists was not significant, $t(23) = 0.15$.

In contrast to the results of the previous experiment, the 6-month-olds did not show any tendency to listen longer to the strong/ weak lists than to the weak/strong lists. So it does not appear to be the case that infants simply find the words from the strong/weak lists more interesting to listen to. Rather, it appears that the preference that the 9-month-olds demonstrated in the previous experiment may come about as a result of their increasing familiarity with the sound patterns of English. To confirm this developmental change, we submitted the data from the 6- and 9-month-olds to an ANOVA of a 2 (age) × 2 (stress pattern) design. Significant main effects were observed for both age, $F(1, 92) = 10.69$, $p < .003$, and stress pattern, $F(1, 92) = 7.01$, $p = .01$. Most important, there was a significant interaction between these two variables, $F(1, 92) = 7.13$, $p < .01$. This latter result supports the notion of a developmental trend in the infants' tendency to listen significantly to words following the strong/weak, as opposed to the weak/ strong, stress pattern.

Thus, the results of these first two experiments suggest that some time between 6 and 9 months of age, American infants begin to show differences in their listening times to lists of words that do or do not follow the predominant stress pattern in English. Our interpretation of these results assumes that it is the stress patterns of the words that infants are responding to. However, to this point, we have not ruled out the possibility that the words from the two types of lists could also embody other commonalities that have to do with their phonetic and phonotactic structure. Indeed, Jusczyk et al. (1993) have shown that 9-month-old infants are sensitive to phonetic and phonotactic properties of words in their native language. Consequently, it is possible that infants in the present study were responding to these features of the words rather than to the stress patterns. For this reason, we decided to conduct the following experiment.

Experiment 3

One means of determining whether infants were responding to the stress patterns of the words in the lists or to other features having to do with their phonetic and phonotactic structure is to eliminate, or at least

greatly reduce, the availability of the latter types of cues in the input. Low-pass filtering the input at a suitable level is one means of achieving this objective that has been used in previous studies with infants (e.g., Cooper & Aslin, 1990; Fernald, 1989; Jusczyk et al., 1992, 1993; Mehler et al., 1988). Thus, in the present experiment, 9-month-old American infants were tested on low-pass filtered versions of the strong/weak and weak/strong lists. If infants are simply responding to the phonetic and phonotactic features of the words from the lists, then one would expect that the preference for the strong/weak lists would disappear under conditions of low-pass filtering. Alternatively, if infants are responding to the stress patterns of the words in the lists, then the preference for the strong/weak lists should remain intact even when the lists are low-pass filtered.

Method

Subjects.—Twenty-four infants of approximately 9 months of age (15 males and nine females) from the suburban Buffalo area were tested. The infants had an average age of 40 weeks, 4 days (range: 38 weeks, 3 days to 43 weeks, 2 days). Five additional infants were tested but excluded for the following reasons: parent failed to keep the infant centered during the experiment (3), looking times for each side were under 3 sec (1), experimenter error (1).

Stimulus materials.—The same lists were used as in the previous two experiments. However, after the conversion of digitized files to audio output by the D/A converter, the resulting signals were passed through a Krohn-Hite filter with the low-pass cutoff set to 400 Hz with an attenuation slope of 48 dB per octave. This filter level was sufficient to eliminate most of the distinctive phonetic information from the samples while leaving intact prosodic features, such as intonation, stress, and rhythm. The filtered samples were output to the amplifier and then to the loudspeaker in the testing room. Note that low-pass filtering does tend to reduce the overall amplitude of the speech sounds. However, the loudness levels on the amplifiers were adjusted to ensure that the lists were played at the same sound pressure levels as in the previous experiments (i.e., 72 ± 2 dB [C] SPL).

Apparatus and procedure.—Slight modifications were made to the apparatus so that the test sessions could be videotaped as well as scored on line by the observer in the test room. A hole of about 8 cm in diameter was cut into the pegboard of the center panel 8 cm below the green flashing light. A JVC compact video camera (GR-303) was aligned with the hole behind the pegboard. The camera was used with the existing lighting in the room. In all other respects, the apparatus and procedure remained the same as in the previous experiments.

Results and Discussion

The amount of time that each infant oriented to the loudspeaker on each trial was recorded. The average looking times were 7.37 sec (SD = 2.30 sec) for the weak/strong lists and 8.25 sec (SD = 2.23 sec) for the strong/weak lists. Fifteen of the 24 infants showed longer looking times for the strong/weak lists. A paired t test confirmed that the difference in orientation times to the strong/weak and weak/strong lists was significant, $t(23) = 2.05$, $p = .05$. Thus, the preference for the lists of words with strong/weak stress patterns remained even though the samples were low-pass filtered. This suggests that the prosodic differences between the two types of lists, which remain after low-pass filtering, are sufficient to produce longer listening times to the strong/weak words. Note that in other experiments, low-pass filtering eliminated preferences that are apparently based on the phonetic and phonotactic properties of words (Jusczyk et al., 1993).

The looking times in this experiment appeared to be longer than in Experiment 1, which used the unfiltered versions of the word lists. To examine this possibility we submitted the data from both experiments to an ANOVA of a 2 (experiment) × 2 (list type) design. Both the main effects of experiment, $F(1, 92) = 11.92$, $p < .001$, and list type proved to be significant, $F(1, 92) = 10.72$, $p < .002$. However, there was no indication of a significant interaction between these two factors, $F(1, 92) = 1.85$, $p > .20$. Hence there is some indication that the infants actually listened longer overall to the low-pass filtered stimuli than they did to the unfiltered versions. Perhaps the reduction of phonetic information rendered the low-pass filtered versions more novel for infants. This is not the pattern typically observed in previous studies using low-pass filtering (e.g., Jusczyk et al., 1992, 1993), where, if anything, listening times to low-pass filtered stimuli tended to be shorter than for unfiltered stimuli. We suspect that the present difference is simply due to differences between the two subject groups, and perhaps to nothing more than the fact that the subjects in Experiment

3 were just a little older than the ones in Experiment 1.

The use of the videocamera throughout the sessions also provided us with the opportunity to conduct reliability checks on the times recorded by the observer during testing. The original observer and a second observer independently scored videotapes with the soundtracks turned off for 22 of the 24 subjects in the study (unfortunately, the data from the other two subjects could not be scored on video because of errors in recording the sessions). We conducted three different types of reliability checks. The first involved a correlation of the difference in looking times for the strong/weak and weak/strong lists for each infant as reported by the same observer when viewing each infant "live" and on videotape. The Pearson product-moment correlation indicated high agreement across these observations, $r(20)$ = .924. Next, we examined agreement between the "live" observations and those made from the videotapes by the second observer. Once again, there was excellent agreement, $r(20)$ = .937. Finally, there was excellent agreement between the looking times recorded by both observers while viewing the videotapes, $r(20)$ = .961. Thus, there was high agreement between different observers of the same infants. As a further check on observer agreement, we took a closer look at possible discrepancies between the live and videotape looking time judgments by looking at differences in recorded times on a trial-by-trial basis. On 189 trials (72%), the discrepancy between the recorded time of the live observer and a different observer viewing the videotape was less than 0.5 sec. On 41 trials (15%), there was a discrepancy of 1 sec or more. For these discrepancies of 1 sec or more, we examined the pattern of differences between the two observers to see whether there was any systematic tendency for the live observer to overestimate looking times on the strong/weak trials and to underestimate them for the weak/strong trials. A t test revealed that there was no systematic difference between observers across these trials, $t(39)$ = 0.56, p = .60. Therefore, the high agreement be-

tween the times recorded during the "live" observations and those of the videos without the soundtracks is an indication that the experimenter's judgments of looking times were not unwittingly biased by the possibility that she may have heard a portion of the soundtrack despite the masking noise.

The present results with low-pass filtered versions of the lists suggest that 9-month-olds do respond to differences in the stress patterns of the words. Thus, even when the availability of potential phonetic and phonotactic cues in the input is greatly reduced, the infants still listen longer to the lists with words that follow a strong/weak stress pattern.

General Discussion

The present study has shown that during the latter half of the first year of life infants are learning about characteristic patterns in the sound structure of their native language. In particular, the results of the first experiment demonstrated that 9-month-old American infants listened longer to lists of items that conform to the predominant strong/weak stress pattern of English words than to lists that did not display this pattern. The results of the second experiment indicated that 6-month-old American infants did not exhibit the same preferences for words with strong/weak stress patterns as did the older infants, despite the fact that other experiments with the same looking paradigm have demonstrated that 6-month-olds are capable of many speech discriminations.[2] Consequently, it appears that it is not simply the case that, in general, infants find words with a strong/weak pattern more interesting to listen to than ones following a weak/strong pattern. Rather, it seems that the tendency of 9-month-olds to listen longer to words with the strong/weak pattern can be attributed to their increasing familiarity with the predominant stress pattern of English. This interpretation of the 9-month-olds' behavior receives additional support from the results of the third experiment. Low-pass filtering the lists left the stress patterns of the words intact while removing most of their phonetic

---

[2] One of the reviewers suggested the possibility that 6-month-olds might have shown a preference for the Strong/Weak lists had we used infant-directed instead of adult-directed speech. This is certainly an intriguing possibility. However, recall that prosodic distinctions were more marked in our Weak/Strong lists than in our Strong/Weak lists—that is, our Weak/Strong lists were actually closer to the prosodic exaggeration typical of child-directed speech. By using adult-directed speech and isolated words, we believe that we have made the strongest possible test of the hypothesis that prelinguistic infants are sensitive to stress pattern information about individual words.

content. Nevertheless, once again infants listened longer to lists composed of words with strong/weak stress patterns.

Thus, the present results suggest that the characteristic stress pattern of English words is a very salient feature for American infants at around 9 months of age. It is worth recalling that the stress patterns of words on both types of lists are legal patterns within the English language, but that the strong/weak pattern is far more prevalent in the input. The 9-month-olds therefore appear to be distinguishing between frequently and less frequently occurring prosodic patterns within the language. This is a relatively fine-grained prosodic distinction. One might expect that preferences based on coarser-grained prosodic distinctions might be developed even earlier. Indeed, there is evidence that this is the case. Mehler et al. (1988) found that newborn French infants displayed preferences for utterances in their native language as opposed to ones in a foreign language, and that this preference occurred even when the utterances were low-pass filtered. Moreover, they also found that 2-month-old American infants distinguished between English and Italian utterances, even when these utterances were low-pass filtered. Thus, by 2 months of age, American infants display some ability to distinguish between English and foreign language utterances on the basis of the prosodic information that is left intact after low-pass filtering.

Of course, the samples that Mehler et al. used were utterances from continuous speech. Consequently, they provided a far richer source for distinguishing native from foreign language utterances than one might expect from the information available in the stress patterns of individual words. Nevertheless, there are indications that, by 6 months of age, American infants are sensitive to prosodic differences between English and some foreign languages at the word level. Jusczyk et al. (1993) found that American 6-month-olds listened significantly longer to lists of unfamiliar English words than to Norwegian words produced by the same talker. Furthermore, the longer listening times for the English words were still present when the lists were low-pass filtered. Norwegian word prosody differs from English in several ways. First, in English, stressed syllables tend to have more pitch movements, greater amplitude, and increased durations relative to unstressed syllables (Crystal & House, 1988; Ladefoged, 1975). By comparison, in Norwegian, these

three prosodic dimensions exhibit a different relation to syllable stress. For example, Norwegian words characteristically have a pitch rise on the word final syllable, and stressed syllables tend to have low pitch (Haugen & Joos, 1972). Such differences in overall prosodic structure can lead to preferences for the native language structure at 6 months, as Jusczyk et al.'s (1993) results show.

However, not all foreign languages differ from English in characteristic word prosody. For instance, Dutch resembles English in that it is a stress language, that is, it displays an opposition between strong and weak syllables (Rietveld & Koopmans-van Beinum, 1987). Furthermore, the pitch, amplitude, and durational correlates of stress in Dutch tend to pattern the same way as in English (Rietveld, 1988). Therefore, it was not surprising that Jusczyk et al. found that the 6-month-olds showed no preference for English words when they were pitted against Dutch words. One aspect of the native language that may be particularly salient for infants during the early phases of acquisition is its characteristic prosodic structure; during the latter half of the first year of life, infants gradually accumulate information about this aspect of their language.

In addition to infants' increased attention to stress patterns of the native language, sensitivity is of course also developing to other structural features of the input. Thus Kuhl et al. (1992) have recently reported that infants' categorization of vowels begins at 6 months to organize around the prototypical values characteristic of the native language. Similarly, infants at this age show signs of becoming more attuned to prosodic markers of important perceptual units in the input, such as clauses (Hirsh-Pasek et al., 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Wright Cassidy, 1989) and phrases (Jusczyk et al., 1992). Perhaps, then, it is not surprising that increases in sensitivity that infants display toward characteristics of the native language structure also appear to go hand in hand with a decline in sensitivity toward certain features that do not regularly appear in the language. For example, Werker and her colleagues (Werker & Lalonde, 1988; Werker & Tees, 1984) have documented a decline in sensitivity during the latter half of the first year of life by infants from English-speaking homes to certain phonetic contrasts not found in English (but see also Best, McRoberts, & Sithole, 1988, for indica-

tions that not all foreign language contrasts undergo the same decline). Apparently, during the latter portion of the first year, infants' attention is becoming more closely focused on characteristics that are particular to the native language that they are acquiring. In this respect, it would be interesting to see how infants acquiring a different native language with very different prosodic structure, such as French, Norwegian, or Japanese, might respond to the English word lists.

It is precisely around 9 months of age that one would expect that infants might be beginning the development of a lexicon in the native language. To succeed in this task, they need to have some means of recovering the acoustic structure of lexical items from continuous speech. We have suggested that prosodic structure may be what infants rely upon to solve this initial segmentation problem. Thus we would account for our finding that infants are sensitive to characteristic prosodic patterns in the lexical inventory of English by suggesting that this sensitivity may have originated in the prelinguistic infants' use of prosodic structure as a way of solving the segmentation problem posed by the fact that the words that they seek to learn occur not in isolation but embedded in continuous speech.

There is a sense in which continuous speech poses the same segmentation problem to all listeners, whether or not they are already in possession of a lexicon. It has been proposed, however, that possession of a lexicon can obviate the problem for adult listeners, in that as any word is recognized, it will be obvious where it ends and by implication where the next word in the input begins (see Cole & Jakimik, 1978, for an explicit statement of this proposal). In fact, the structure of the English lexicon and the patterns of occurrence of word types in English speech render it unlikely that this type of "segmentation by default" would work efficiently for the listener: the majority of words in English speech are monosyllabic (Cutler & Carter, 1987), and monosyllabic words cannot usually be uniquely identified before their final segment (Luce, 1986) and often, in fact, are not identified until after their acoustic offset (Bard, Shillcock, & Altmann, 1988; Grosjean, 1985). Thus, it is perhaps not surprising to find that "segmentation by default" is apparently not the option chosen by English listeners; instead, there is evidence that English listeners use explicit segmentation procedures in recognizing words in continuous speech. Moreover, the seg-

mentation procedures are based on precisely those characteristics of the English vocabulary on which the present study has focused, namely, the distinction between strong and weak syllables, and the asymmetric distribution of stress patterns within the vocabulary. Cutler and Norris (1988) showed that the detection of real words embedded in nonsense strings was inhibited if the embedded word formed part of two strong syllables, as opposed to a strong and a following weak syllable; they argued that listeners segment speech at strong syllable onsets, and that detecting an embedded word that forms part of two strong syllables is difficult because it requires recombination of speech material across a segmentation point. The procedure of segmenting continuous speech at strong syllable onsets, which Cutler (1990) called the metrical segmentation strategy, should, given the asymmetries in the vocabulary, work very efficiently to locate actual lexical word onsets in continuous speech input. Further evidence that English listeners use this procedure was provided by Cutler and Butterfield (1992) in a study of spontaneous and laboratory-induced misperceptions of continuous speech. Where word boundaries were misperceived, listeners' errors were systematically related to the prosodic structure of the input: word boundaries were inserted prior to strong syllables, but deleted prior to weak syllables. In other words, listeners were apparently acting according to the predictions of the metrical segmentation strategy by assuming that any strong syllable was most likely to be the initial syllable of a lexical word.

The prosodic asymmetries in the English vocabulary are thus exploited by listeners at all stages of development. Adults use prosodic probabilities as the basis for segmentation procedures that improve the efficiency of word recognition in continuous speech. Infants at the stage at which a lexicon is being acquired already show sensitivity to prosodic probabilities, in that, as the present studies have shown, they demonstrate a preference for more likely stress patterns over less likely patterns. We have argued that this sensitivity could reflect infants' use of prosodic structure as a means of beginning the task of segmenting continuous speech into words, which, in its turn, is a prerequisite for the development of a lexicon. Prosodic structure is one of the most salient features of a language's lexical repertoire, and it is also one of the earliest features to which language users are sensitive.

Appendix

List of matched stimulus words used in the samples

	Weak/Strong	*Strong/Weak*
1.	abeam	beaver
2.	provide	idle
3.	ablaze	lazy
4.	abloom	loony
5.	aboard	orbit
6.	bizarre	barber
7.	supreme	reamer
8.	comply	pliant
9.	befall	falter
10.	upon	ponder
11.	depart	partner
12.	derive	hydrant
13.	aback	blacken
14.	comport	comet
15.	condone	donor
16.	perceive	even
17.	apace	pagan
18.	facade	solder
19.	denote	notion
20.	abut	butter
21.	commode	motor
22.	contort	torture
23.	devise	visor
24.	elude	looter
25.	gavotte	vodka
26.	across	crossing
27.	approach	prosy
28.	amaze	matron
29.	cerise	rebus
30.	define	final
31.	macaw	cautious
32.	arraign	rainy
33.	escape	patron
34.	inflate	fated
35.	pomade	neighbor
36.	assume	humor
37.	acquire	wireless
38.	deceit	teeter
39.	alone	loner
40.	begrime	rhymer
41.	bewail	whaler
42.	ascribe	sprightly
43.	restore	stalwart
44.	bespeak	speaker
45.	charade	ardour
46.	abode	bogus
47.	report	porter
48.	remain	manger
49.	between	teeny
50.	abound	bounden
51.	afoul	fowler
52.	aground	brownie
53.	arouse	rousing
54.	abide	spider
55.	affright	writer
56.	align	liner
57.	apprize	prizes
58.	assign	sinus
59.	condign	diamond
60.	contrive	private
61.	demise	miser
62.	polite	lighter
63.	enlarge	sergeant
64.	garotte	rocket
65.	caprice	rhesus
66.	ravine	venous
67.	aloof	lucid
68.	pollute	flutist
69.	attune	tuna
70.	cheroot	rooting
71.	awake	sacred
72.	savant	bondage
73.	before	former
74.	perturb	turban
75.	advance	rancid
76.	attempt	empty
77.	observe	perfect
78.	devout	outer
79.	refute	feudal
80.	occur	curdle
81.	engulf	sulphur
82.	attend	ending
83.	refer	further
84.	repel	elder
85.	regard	gargle
86.	return	turnip
87.	demand	mantle
88.	accost	coster
89.	resent	gentle
90.	distilled	builder
91.	avert	virgin
92.	corrode	rover
93.	offence	enter
94.	deserve	servant
95.	remark	argus
96.	decay	cable

References

Bard, E. G., Shillcock, R. C., & Altmann, G. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effect of subsequent context. *Perception and Psychophysics*, **44**, 395–408.

Best, C. T. (1991, April). *Phonetic influences on the perception of non-native speech contrasts by 6–8- and 10–12-month-olds*. Paper presented at biennial meeting of the Society for Research in Child Development, Seattle.

Best, C. T., McRoberts, G. W., Goodell, E., Womer, J., Insabella, G., Klatt, L., Luke, S., & Silver, J. (1990, April). *Infant and adult perception of non-native speech contrasts differing in relation to the listeners' native phonology*. Paper presented at meeting of the International Conference on Infant Studies, Montreal.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 245–360.

Carlson, R., Elenius, K., Granstrom, B., & Hunnicutt, S. (1985). Phonetic and orthographic properties of the basic vocabulary of five European languages. In *Quarterly progress and status report* (Vol. 1, pp. 63–94). Stockholm: Speech Transmission Laboratory.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1993). *Do infants hear word boundaries? An empirical approach to the bootstrapping problem for lexical acquisition.* Manuscript in preparation.

Cole, R. A., & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing.* New York: Academic Press.

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior & Development,* 4, 219–223.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development,* 61, 1584–1595.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected speech signals: Syllabic stress. *Journal of the Acoustical Society of America,* 83, 1574–1585.

Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 105–121). Cambridge, MA: MIT Press.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language,* 31, 218–236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language,* 2, 133–142.

Cutler, A., & Norris, D. G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance,* 14, 113–121.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development,* 8, 181–195.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development,* 10, 279–293.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mother's speech to newborns. *Developmental Psychology,* 20, 104–113.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B. D., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language,* 16, 477–501.

Friedlander, B. Z., & Wisdom, S. S. (1971). *Preverbal infants' selective operant responses for different levels of auditory complexity and language redundancy.* Paper presented at the annual general meeting of the Eastern Psychological Association, New York.

Gleitman, L., Gleitman, H., Landau, B., & Wanner, E. (1988). Where learning begins: Initial representations for language learning. In F. Newmeyer (Ed.), *The Cambridge linguistic survey.* Cambridge, MA: Harvard University Press.

Glenn, S. M., Cunningham, C. C., & Joyce, P. F. (1981). A study of auditory preferences in nonhandicapped infants and infants with Down's syndrome. *Child Development,* 52, 1303–1307.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics,* 38, 299–310.

Harrington, J., & Johnstone, A. (1987). The effects of equivalence classes on parsing phonemes into words in continuous speech recognition. *Computer Speech and Language,* 2, 273–288.

Haugen, E., & Joos, M. (1972). Tone and intonation in East Norwegian. In D. Bolinger (Ed.), *Intonation* (pp. 414–436). Baltimore: Penguin.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition,* 26, 269–286.

Jusczyk, P. W. (1989, April). *Perception of cues to clausal units in native and non-native languages.* Paper presented at the biennial meeting of the Society for Research in Child Development, Kansas City, MO.

Jusczyk, P. W., Friederici, A. D., Wessels, J., & Svenkerud, V., & Jusczyk, A. M. (1993). *Infants' recognition of foreign versus native language words.* Manuscript submitted for publication.

Jusczyk, P. W., Kemler Nelson, D. G., Hirsh-Pasek, K., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology,* 24, 252–293.

Jusczyk, P. W., & Thompson, E. J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception & Psychophysics,* 23, 105–109.

Karzon, R. G. (1985). Discrimination of a polysyllabic sequence by one- to four-month-old infants. *Journal of Experimental Child Psychology,* 39, 326–342.

Kelly, M. H., & Martin, S. (in press). Sensitivity to probabilistic information: A basic component to perception and cognition. *Lingua.*

Kemler-Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Wright-Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language,* 16, 55–68.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169–226). Cambridge, MA: MIT Press.

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science,* 1, 70–73.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science,* 255, 606–608.

Ladefoged, P. (1975). *A course in phonetics* (pp. 217–224). New York: Harcourt Brace Jovanovich.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception and Psychophysics,* 39, 155–158.

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception,* 7, 491–497.

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 236–262). Hillsdale, NJ: Erlbaum.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition,* 29, 144–178.

Miller, G. A. (1964). The psycholinguists. *Encounter,* 23, 29–37.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (in press). Preferences for infant-directed over adult-directed speech: Evidence from 7 week old infants. *Infant Behavior and Development.*

Rietveld, A. C. M. (1988). Woordklemtoon in spraak. In M. P. R. van den Broeke (Ed.), *Ter sprake* (pp. 132–143). Dordrecht: Foris.

Rietveld, A. C. M., & Koopmans-van Beinum, F. J. (1987). Vowel reduction and stress. *Speech Communication,* 6, 217–229.

Spring, D. R., & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research,* 20, 224–232.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology,* 24, 672–683.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development,* 7, 49–63.