Perception & Psychophysics
1993, 54 (3), 287-295

Phonotactic knowledge of word boundaries and its use in infant speech perception

ANGELA D. FRIEDERICI
Free University Berlin, Berlin, Germany

and

JEANINE M. I. WESSELS
Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

The development of a lexicon critically depends on the infant's ability to identify wordlike units in the auditory speech input. The present study investigated at what age infants become sensitive to language-specific phonotactic features that signal word boundaries and to what extent they are able to use this knowledge to segment speech input. Experiment 1 showed that infants at the age of 9 months were sensitive to the phonotactic structure of word boundaries when wordlike units were presented in isolation. Experiments 2 to 5 demonstrated that this sensitivity was present even when critical items were presented in context, although only under certain conditions. Preferences for legal over illegal word boundary clusters were found when critical items were embedded in two identical syllables, keeping language processing requirements and attentional requirements low. Experiment 6 replicated the findings of Experiment 1. Experiment 7 was a low-pass-filtered version of Experiment 6 that left the prosody of the stimulus items intact while removing most of the distinctive phonotactic cues. As expected, no listening preference for legal over illegal word boundary clusters was found in this experiment. This clearly suggests that the preferential patterns observed can be attributed to the infants' sensitivity to phonotactic constraints on word boundaries in a given language and not to suprasegmental cues.

The understanding of spoken language requires the identification of individual words in the utterance. The adult language user's speech segmentation process is supported by acquired lexical knowledge. The prelinguistic child, however, cannot rely on lexical knowledge and must, therefore, base segmentation of continuous speech initially on cues which are nonlexical but which in the end will lead to correct segmentation into lexical elements. The difficult task of identifying words in the speech stream must have been solved—at least partly—during the first 14-16 months of life, since it is around that time that young children start to produce one-word utterances. To master this clearly, some kind of segmentation process must have been applied, inasmuch as the child's language input does not consist solely of isolated words (Mehler, Dupoux, & Segui, 1990). But how does the young child achieve segmentation in the absence of lexical knowledge?

One possibility is that this achievement is supported by some innate mechanism. However, although infants come into the world equipped with a certain attraction toward human language (Colombo & Bundy, 1981; Glenn, Cunningham, & Joyce, 1981), in particular, when it is realized in an infant-directed form (Cooper & Aslin, 1990; Fernald, 1985; Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978; Pegg, Werker, & McLeod, in press), it is not likely that infants can use innate knowledge about the possible structure of words since these differ greatly from language to language. This means that the child must acquire this knowledge by extracting it from nonlexical cues given in the language input.

There are basically two types of cues the young child might use for the initial segmentation of speech and the identification of words: suprasegmental cues, such as prosodic features, and segmental cues, such as phonotactic features. With regard to suprasegmental cues, it has been shown that prosodic features are indeed used by the infant to structure language input. Newborns, in only a few days after birth, already demonstrate sensitivity to the overall intonational characteristics of their native language (Mehler et al., 1988). At the age of 6 months, they show sensitivity to the prosodic characteristics of clause boundaries (Hirsh-Pasek et al., 1987), and at the age of 9 months, they are sensitive to the prosodic features of phrase boundaries (Jusczyk et al., 1992). According to Kemler Nelson (1989), they do not appear to be sensitive to word boundaries until the age of 11 months. However, Jusczyk, Cutler, and Redanz (in press) provide evidence suggesting that infants are sensitive to the prosodic structure of words at the age of 9 months. They report that

This research was supported by a grant from the Deutsche Forschungsgemeinschaft (Fr 519/9-1) and by the Alfried Krupp von Bohlen und Halbach Science Prize awarded to A.F. Special thanks are extended to Peter Jusczyk, who opened the door to infant research for us, and to the Max-Planck Institute for Psycholinguistics in Nijmegen, without whose support this work would not have been possible. Correspondence may be addressed to A. D. Friederici, Institute für Psychologie, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.

9-month-old, but not 6-month-old, English-learning infants listened longer to English bisyllabic words with the (in English) more frequent strong/weak stress pattern than to those with the less frequent weak/strong stress pattern when words were presented in isolation.

Another possible class of cues infants might use to support the speech segmentation process involves phonotactic features. Young infants have been shown to be able to discriminate nearly every phonetic contrast (for reviews, see Eimas, Miller, & Juscyzk, 1987; Kuhl, 1987; Werker & Pegg, 1992). Moreover, work from different laboratories indicates that infants' phonetic sensitivity changes as a function of particular language inputs (e.g., Aslin, Pisoni, Hennessy, & Perey, 1981; Werker & Tees, 1984). Sensitivity to language-specific vowel quality appears to be observable at the age of 6 months (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), and sensitivity to language-specific consonant phonetic and phonotactic quality can be observed in 9-month-old, but not in 6-month-old infants (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, in press).

Jusczyk, Friederici, et al. (in press) provide evidence that infants at the age of 9 months prefer to listen to words that contain consonants that are part of their native language's inventory rather than to those that belong to the inventory of a foreign language. Moreover, they even appear to be sensitive to language-specific phonotactic constraints, since they prefer to listen to words phonotactically legal in their native language (e.g., *knevel*, in Dutch) rather than to words that contain consonant clusters belonging to a foreign but not to their native language (e.g., *dwell*, in English). In the latter case, consonants *d* and *w* both belong to the phonetic inventory of Dutch, but they never appear in the *dw* ordering, an illegal combination in Dutch.

Thus, it seems that during the second half of the first year of life infants have developed some of the potential prerequisites necessary to identify words in the speech stream. It is an open question, however, as to whether infants actually use this knowledge to segment the auditory speech stream into words. So far, sensitivity to word structure and language-specific phonotactic knowledge in 9-month-olds has been demonstrated only when words were presented in isolation (Jusczyk, Cutler, & Redanz, in press; Jusczyk, Friederici, et al., in press).

Moreover, it is not clear how the infant manages the first step in acquiring word-specific knowledge. The possession of knowledge concerning language-specific prosodic word structure or phonotactic word structure seems to presuppose the identification of wordlike units as a necessary condition. The most simple way to achieve this condition would be to receive isolated words as input. This, however, is not what infants usually receive as input from their parents (Mehler et al., 1990), although parents may use other ways of supporting the infants' search for word boundaries. They may choose to place critical words at the beginning or the end of an utterance. Whether parents indeed use such a strategy was investigated recently by Aslin (1993), who analyzed infant-directed speech in a search for possible cues that the mothers' inputs might provide for word segmentation. He gathered samples of maternal speech directed to 12-month-olds, with the mother (1) teaching the infant a new word and (2) reading from a story book. The maternal speech was collected from two languages, English and Turkish, the latter being a verb-final language. He found that mothers rarely produced target words in isolation, but rather highlighted the target word by exaggerating pitch contours and by placing the target in utterance-final or utterance-initial position. Target nouns were even placed in utterance-final position in Turkish, in which this construction is clearly ungrammatical. However, by placing target words in such a salient position, mothers reduce the segmentation problem for the child. It is not unlikely that the infant's initial search for word boundaries is supported by these aspects of infant-directed speech.

In the present study, we did not follow up on the question of how infant-directed speech might support initial steps in the process of identifying word boundaries. Rather, we supposed that aspects like those investigated by Aslin (1993) do plausibly help in the infant's initial identification of word boundaries. We hypothesized that once regularities about word boundaries are extracted from experienced linguistic input, infants might use this knowledge as a basis for the segmentation of novel speech input.

We will report a series of experiments that investigated infants' knowledge about phonotactic regularities of possible word boundaries as well as the infants' ability to use this knowledge for speech segmentation. In the first experiment, we sought to determine the age at which infants become sensitive to phonotactic constraints about possible word boundaries in their native language. For example, certain sequences of phonetic elements are legal only at the beginning of a word (e.g., str-, as in *street*) but not at its end (e.g., -str). In Experiments 2 through 5, we investigated whether, and under what circumstances, infants are able to use this knowledge to identify possible word boundaries in connected speech. In Experiment 2, one-syllable wordlike units with phonotactically legal and illegal onset and offset clusters were presented embedded in a two-syllable context consisting of two nonidentical syllables. As no preferential pattern was found in this experiment, in Experiment 3 the language processing requirements were reduced by simplifying the context to two identical context syllables. Since no preference was observed in this experiment either, in Experiment 4 the perceptual requirements were reduced by using an infant-directed-speech mode. As still no preference was found in this experiment, in Experiment 5 the attentional requirements were reduced by shortening the interstimulus interval between the stimuli within the trials. In this experiment, infants showed a significant preference for items with legal over illegal word-offset clusters. Experiment 6 served as a replication of Experiment 1 with an interstimulus interval similar to that of Experiment 5. Finally, in Experiment 7 we ran a low-pass-filtered version of Experiment 6 in order to demonstrate that the observed sensitivity to phonotactically marked word boundaries was

indeed due to the infants' sensitivity to phonotactic rather than to suprasegmental features.

EXPERIMENT 1

In Experiment 1, we investigated infants' sensitivity to phonotactically legal and illegal word onset and offset clusters when wordlike units were presented in isolation, that is, in word lists with pauses between each wordlike unit.

Previous work had shown that infants between the ages of 6 and 10 months demonstrate a language-specific reorganization in phonetic perception (e.g., Werker & Tees, 1984; Werker & Lalonde, 1988). A study by Jusczyk, Friederici, et al. (in press) indicates that 9-month-old, but not 6-month-old, infants demonstrate sensitivity to phonotactic features of their native language. The phonotactic clusters tested in that study consisted of phonemes that were legal in the infant's native language and in a foreign language, but whose ordering was legal in one of them but illegal in the other. Sound-structure features of words, that is, the stress pattern of words, have been shown to affect infants' language perception at the age of 6 months (Jusczyk, Cutler, & Redanz, in press) and even at 4.5 months (Jusczyk, 1989). We therefore tested a group of 9-month-old, a group of 6-month-old, and a group of 4.5-month-old Dutch infants to determine at what age they were sensitive to particular phonotactic features of word boundaries in Dutch.

Method

Subjects. Three groups of infants, 9, 6, and 4.5 months old, each age group consisting of 24 infants, participated in this experiment. All infants came from monolingual Dutch-speaking families; they were healthy and had uncomplicated prenatal and postnatal histories.

Forty-three 4.5-month-olds were examined and 19 were excluded, 8 due to lack of interest, 1 because of crying, 3 because of restlessness, 6 because of looking times that were too short, that is, under 3 sec, and 1 because of too many missing data, that is, more than one trial had a looking time shorter than 1 sec. The remaining 24 subjects, with a mean age of 4.6 months (range 4.3 to 4.9), were used in the analysis. The 6-month-old infants averaged 6.2 months in age (range 5.9 to 6.5). To obtain the necessary number of subjects, 30 infants were tested; 6 were excluded from the analysis, 2 for crying, 1 for failing to look for longer than 3 sec, 1 for experimental failure, 1 because of a cold, and 1 because of experimenter error. The 9-month-old infants had a mean age of 9.1 months (range 8.6 to 9.4). Thirty-two subjects were tested and 8 were excluded, 2 for crying, 5 for restlessness, and 1 because it was afraid.

Stimulus material. All the stimuli were generated from a set of recordings of the same female voice. The speaker understood the purpose of the research and accordingly produced the speech samples very accurately, as the experimental restrictions required. The speech samples consisted of one-syllable items that were either phonotactically legal or phonotactically illegal. Each speech trial contained 15 items with an interstimulus interval of 1,250 msec between the items. There were 12 test speech trials. The length of the speech trials was kept constant (mean legal = 22 sec 408 msec, mean illegal = 22 sec 430 msec). To avoid familiarity or recognition of the legal items presented, only nonwords were used. To create illegal strings, the following procedure was used: A legal onset cluster was inserted at the end of the item leading to an illegal phonotactic word offset, for example, /BRef/ vs. /feBR/ (critical cluster printed in capital letters). Illegal onset structures were created by inserting a legal offset cluster at the beginning of the word, for example, /muRT/ vs. /RTum/. Legal onset and offset clusters that matched in frequency were selected using the information database CELEX (1990), a Dutch corpus containing approximately 42 million entries allowing for the selection of words or syllables according to different criteria, for example, phonemic, phonotactic, or morphological criteria, as well as frequency. To identify the critical legal and illegal clusters, the CELEX database was searched for the selection of all one-, two-, and three-syllable words with (1) a two-consonant word onset and (2) a two-consonant word offset. For each of these two categories, the mean frequency of words that were members of a particular onset class was calculated. A total of 35 critical onset clusters and a total of 30 critical offset clusters were identified. These clusters were categorized in terms of low, medium, or high frequency (the mean frequency was calculated over one-, two-, and three-syllable words). Only clusters from the category of medium frequency were selected as experimental items (mean frequency of 21-30 per million). This selection procedure resulted in 12 critical onset clusters (BR, GL, KL, KR, SG, SGR, SL, SN, STR, TR, VL, ZW) and 5 critical offset clusters (KS, RM, RT, GT, NT).

Apparatus. The infant was seated on the caregiver's lap in the center of a three-sided enclosure, 4×6 ft on three sides and open at the back. A green light was mounted at eye level on the center panel. A red light and acoustic speaker were mounted at eye level on each of the side panels, 78° to the left and right of the infant when it was facing midline. Utilizing a 286 microprocessor, the Speech Server, speech output was generated by means of a SORCUS ML4 board, a Z80 microprocessor, and two digital/analog converters (for left and right presentation). A second AT-PC was responsible for running the experiment. This computer configuration was placed behind the central panel. The experimenter was seated behind the computer and watched the infants' behavior on a monitor that was connected with a video camera placed above the green light in the center panel. Information about the infants' head orientation, that is, the timing and direction of the infants' head orientations, was entered into the computer on-line by the experimenter, who was blind to the conditions.

Procedure. The present study used a version of the head-turn procedure, which differed from that introduced by Fernald (1985). In the present version, the legal speech samples were consistently played through a loudspeaker situated to one side of the infant and the illegal ones were consistently played through a loudspeaker situated to the other side. For half the infants, the legal speech samples were assigned to the right side and the illegal ones were assigned to the left; for the other half, the assignments were reversed. Preferences were indexed by monitoring the durations of the infant's headturns toward one of the two loudspeakers over a set of 12 test trials. These 12 test trials were preceded by 4 preexposure trials, designed to acquaint the infant with the assigned positions of the two kinds of stimulus (see also Hirsh-Pasek et al., 1987).

Each trial was begun by the blinking of the green light in order to draw the infant's attention to the center position. When the observer signaled that the infant was oriented toward midline, the center green light was extinguished and the red light above one of the acoustic speakers began to flash. This light indicated that a speech sample was available on that side, provided that the infant made a headturn of at least 30° in the direction of the corresponding speaker. When the observer detected such a headturn, she pressed a button to start the speech trial at the side corresponding to the blinking light. The speech excerpt was continued either to completion or until the observer indicated that the infant failed to maintain the 30° headturn for at least 2 sec, in which event the sample was terminated immediately. For both the preexposure and the test trials, a silent moving puppet at the center entertained the infant in a short intertrial interval. For half of the subjects, the preexposure trials started with a legal sample; for the other half, they started with an illegal sample. In the preexposure trials, each subject heard

two legal and two illegal samples. The test trials consisted of 6 legal and 6 illegal samples, with the subjects not hearing more than two legal or two illegal trials in a row.

The observer, who started the trials and indicated the occurrence and termination of headturns, was blind as to whether the samples were legal or illegal. Both she and the caregiver who held the infant wore headphones over which music was played to prevent them from hearing the speech samples.

Results

The data from the 12 test trials for each subject were used to calculate the average length of orientation to the legal and illegal speech samples. For the 9-month-olds, the critical difference between the legal and the illegal conditions was significant [$t(23) = 2.42$, $p < .02$]. The mean lengths of orientation were 7.7 sec for the legal condition and 6.0 sec for the illegal condition. This preferential pattern was also present at the individual subject level; 17 of the 24 subjects showed a longer orientation time to the legal condition.

Results of the 6-month-olds, however, showed no preferential pattern. The apparent difference between the legal condition (7.5 sec) and the illegal condition (8.7 sec) did not reach significance in the t test [$t(23) = 1.546$, $p = .136$].

For the 4.5-month-olds, no critical difference between the two conditions was found [$t(23) = 6.554, p = .585$]. The mean orientation times were 8.3 sec for the phonotactic legal condition and 7.6 sec for the phonotactic illegal condition.

Discussion

The results from this first experiment clearly showed that 9-month-olds are sensitive to the phonotactic structure of words of their native language. The younger infants aged 4.5 and 6 months demonstrated no significant differences between the legal and the illegal conditions. A clear preference for the phonotactically legal over the phonotactically illegal word onset and offset structures was observed only at the age of 9 months. At that age, infants seem to possess subtle language-specific phonotactic knowledge about possible and impossible word onsets and word offsets.

The obvious question is whether infants at that age are able to use this knowledge about possible word boundaries to segment connected speech. The potential knowledge about possible word beginnings or endings, which infants may have learned from motherese (Aslin, 1993), could, in principle, be used to segment larger speech samples into wordlike units. The following experiments evaluated this idea.

EXPERIMENT 2

Experiment 2 investigated whether infants were sensitive to the legality of phonotactic word onset and word offset clusters when one-syllable wordlike units containing these clusters were presented in connected speech.

We chose to present the critical one-syllable wordlike units, which contained either legal or illegal onset and offset clusters, in a two-syllable context with one syllable preceding and one syllable following the critical unit, /mig BRef par/ with the critical legal cluster /BR/ at the onset or /wan meGT pun/ with the critical legal cluster /GT/ at the offset. For the legal speech samples, one or more correct segmentations were possible. For example, for the triplet /mif SGir kur/ with the legal onset cluster /SG/, multiple correct segmentations are possible: /mif SGir kur/, /mifS Gir kur/, /mifS Girk ur/. This was true for a number of the legal onset and offset cluster triplets.

In the illegal condition, any segmentation resulted in illegal word boundary clusters. In the construction of the illegal triplets, care was taken that regardless of the particular way segmentation was chosen, infants would always run into the illegal clusters. For example, in the following illegal offset triplet, each of the following segmentations will result in illegal clusters (illegal clusters are underlined):

/mif ri<u>SG</u> kur/ = illegal offset
/mi<u>fr</u> i<u>SG</u> kur/ = illegal offset at 2 positions
/mi fri<u>SG</u> kur/ = illegal offset
/mif riS <u>Gk</u>ur/ = illegal onset.

Thus, for the illegal condition the triplets cannot be segmented in any way that would result in a legal transition between the context syllables and the critical syllable. Thus, if infants showed a preference for the legal over the illegal onset and offset clusters, this would indicate correct segmentation.

Method

Subjects. In this experiment, 48 9-month-old infants participated, they were assigned to two different groups. One group ($N=24$) was tested for the perception of 12 critical-onset-cluster trials; the other group ($N=24$) was tested for the perception of 12 critical-offset-cluster trials. All infants came from monolingual Dutch-speaking families; they were healthy and had uncomplicated prenatal and postnatal histories.

To obtain the necessary number of subjects, a total of 103 infants were tested. For the onset cluster group, we tested 50 infants; 26 were excluded, 5 due to crying, 12 for failing to look for longer than 3 sec, 2 due to experimental failure, 1 because of a cold, 4 due to lack of interest, 1 because it was afraid, and 1 because of experimenter error. The mean age of the subjects participating in the onset condition was 9.2 months (range 8.2 to 9.6). Fifty-three subjects were tested for the offset cluster group; 29 were excluded, 1 due to crying, 5 due to restlessness, 16 due to too-short looking times, 2 because they were tired, 1 because of video failure, and 2 due to lack of interest. The mean age of the subjects in the offset condition was 9.2 months (range 9.0 to 9.5).

Stimulus material. Each of the two conditions (i.e., the onset and offset cluster conditions) contained 12 speech trials. Each speech trial contained 12 triplets constructed as discussed above. Critical syllables with legal and illegal onset and offset clusters were embedded in two-syllable contexts. To meet the criterion that illegal clusters in context would never result in legal clusters regardless of where they were segmented, most, but not all, critical syllables presented in Experiment 1 were also used in Experiment 2. The interstimulus interval between the triplets was 1,250 msec (as in Experiment 1). For the onset cluster condition, the mean lengths of the speech sam-

ples were 32 sec 196 msec for the legal condition and 32 sec 612 msec for the illegal condition; for the offset cluster condition, the mean lengths of the speech samples were 32 sec 998 msec for the legal condition and 33 sec 473 msec for the illegal condition. In both conditions, the critical syllable within each triplet was stressed to ease identification.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

An analysis of variance with the factors of group (onset cluster group vs. offset cluster group) and condition (legal vs. illegal clusters) revealed a main effect of group with shorter looking times for the offset than for the onset cluster group. This effect, however, did not reach full significance [$F(1,239) = 3.71, p < .07$]. There was no significant effect of phonotactic condition or an interaction.

Separate analyses were also run for each group: The results for the onset cluster group did not show any preference for the legal over illegal onset clusters [$t(23) = 1.573, p = .129$]. The mean looking times were 6.7 sec for the legal phonotactic structures and 7.5 sec for the illegal condition. Ten of the 24 infants preferred the legal structures. For the offset cluster group, there was also no preference for the legal over the illegal clusters [$t(23) = 0.244, p = .809$]. The mean looking times were 5.8 sec for the legal phonotactic structures and 6.0 sec for the illegal condition. Twelve of the 24 infants preferred the legal structures.

Discussion

It seems that 9-month-old infants are not capable of detecting the illegality of phonotactic structures in a contextual environment like the one used in this experiment, regardless of the positions of the critical clusters. The present finding of nonpreference could be due to at least three different factors. First, infants at this age may just not be capable of segmenting the speech stream adequately. In this case, legal triplets could be segmented incorrectly, such that legal strings are perceived in the same way as illegal ones, resulting in a nonpreferential pattern. Second, the infants might not be able to detect the illegality of the clusters in context, due to the fact that the task is linguistically and/or cognitively too demanding. The high dropout rates encountered in this experiment indeed suggest that either the task in general was too demanding to hold the infants' attention or that the particular linguistic context used was too difficult to meet the infants' perceptual abilities. To test the latter possibility, the complexity of the linguistic context was reduced in Experiment 3. An earlier study by Goodsitt, Morse, Ver Hoeve, and Cowan (1984) had suggested that infants' recognition performance increased as complexity of the linguistic context dropped. They found that the recognition of a target syllable (e.g., /ba/) was more successful in a redundant context (e.g., /kobako/) than in a nonredundant context (e.g., /kobati/). Similarly, we expected legal and illegal onset and offset clusters to be detected more successfully under a condition in which the target item was embedded in two identical syllables (e.g., /mig BRef mig/) than under the nonredundant context condition used in Experiment 2 (e.g., /mig BRef par/).

EXPERIMENT 3

In Experiment 3, we tested the hypothesis that the linguistic context used in Experiment 2 was too complex. Complexity of the context was reduced by presenting the critical unit in a so-called redundant context, that is, embedded in two identical context syllables.

Method

Subjects. Twenty-four 9-month-old infants (mean age = 9.2 months; range 8.9 to 9.6) from monolingual Dutch-speaking families participated in this experiment; they were healthy and had uncomplicated prenatal and postnatal histories. To obtain the experimental group, 39 subjects were tested. Fifteen were excluded, 3 due to crying, 3 due to restlessness, 7 because their looking times were too short, and 2 because of computer failure.

Stimulus material. The same critical words were used as in the preceding experiments. Instead of using different context syllables, only one syllable (/mig/) was used to precede and follow the critical words. To meet the criteria mentioned in Experiment 2 concerning the legal and illegal transitions between the syllables, too few critical items were available to make up a complete set of 12 onset and 12 offset cluster triplets. In the present experiment, therefore, we used 6 onset and 6 offset triplets in legal and illegal clusters (drawn from Experiment 1). The triplets had the following structure (critical clusters printed in capital letters):

legal condition:	onset: /mig BRef mig/
	offset: /mig diNT mig/
illegal condition:	onset: /mig NTit mig/
	offset: /mig feBR mig/

The mean lengths were 31 sec 812 msec for the legal onset speech trials and 31 sec 103 msec for the legal offset trials. The illegal onset and offset speech trials were 31 sec 478 msec and 31 sec 265 msec, respectively. Onset and offset cluster items were presented in a randomly mixed order. There were 12 critical trials. All other experimental variables were kept constant.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

There was no significant effect of phonotactic structure [$t(23) = 0.882, p = .387$]. Mean duration times collapsed over onset and offset clusters were 6.2 sec for the legal condition and 5.7 sec for the illegal condition. Fourteen of the 24 infants preferred the legal structures.

Discussion

Although mean orientation times of the 9-month-olds demonstrate a pattern in the expected direction, with longer looking times for the legal condition than for the illegal one, the difference is not statistically reliable. A careful intonational analysis screening the materials used in the preceding context experiments showed that the characteristics of the stressed syllables within the triplets more closely resembled an adult-directed speech mode than an infant-directed speech mode. In the speech signals used in this experiment, the vowel lengthening and

exaggerated pitch parameters normally seen in infant-directed speech were not present. An earlier study (Karzon, 1985) had shown that recognition of critical elements in infants' speech perception is eased by an infant-directed speech mode. Thus, we reasoned that the introduction of vowel lengthening and exaggerated pitch in our material might highlight the critical syllable and thereby assist the infants' segmentation.

EXPERIMENT 4

Experiment 4 was conducted to investigate the influence of an infant-directed speech mode upon the infants' ability to detect legal and illegal word boundary clusters in context. To achieve this objective, the characteristics of infant-directed speech were introduced to the speech signal using language material similar to that of Experiment 3. In contrast to Experiment 3, however, the present experiment used only critical offset clusters. This decision was based on the finding that critical items in motherese are most likely to be utterance final (Aslin, 1993). We reasoned that, in addition to the infant-directed speech manipulation, the use of offset clusters might assist the infants' detection of word boundaries.

Method

Subjects. Twenty-four 9-month-old infants (mean age = 9.1 months; range 8.8 to 9.7) from monolingual Dutch-speaking families participated in this experiment; they were healthy and had uncomplicated prenatal and postnatal histories. To obtain the experimental group, 40 subjects were tested; 16 were excluded, 3 for crying, 7 for restlessness, 5 for having too-short looking times, and 1 due to computer failure.

Stimulus material. The stimuli were read by the same female speaker as in the prior experiments. However, the speaker—the mother of a 4-month-old child—was instructed to read the speech samples as if she were reading them to her child. The characteristics of these speech samples showed an increase of decibels, a higher pitch, a lengthening of the critical items, and an overall exaggerated stress pattern. The stimulus material was similar to that of Experiment 3 using redundant context syllables. In contrast to Experiment 3, however, only legal and illegal word-offset cluster triplets were included in the present experiment. The mean lengths were 34 sec 760 msec for the legal speech samples and 35 sec 129 msec for the illegal ones.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

A t test revealed no significant differences between the conditions [$t(23) = 0.019$, $p = .985$]. The mean orientation times were 6.4 sec for the legal condition and 6.5 sec for the illegal condition. Nine of the 24 infant subjects preferred the legal structures.

Discussion

It seems that despite the above-mentioned manipulations, namely, using an infant-directed speech mode and restricting the material to critical word offset clusters, we were not able to observe the expected preference for the legal over the illegal phonotactic word boundaries in connected speech in 9-month-old infants. Instead, the duration times for the two phonotactic conditions observed here are very similar to those observed in the previous experiments. Also, the dropout rates in Experiments 2, 3, and 4 were very high, suggesting that the 9-month-olds' attention may not have been fully captured in these experiments. One possible explanation for the observed rigid nonpreferential pattern may be that the 1,250-msec pauses used as the interstimulus intervals between the triplets were too long, allowing the infants to turn away between the triplets and become distracted. Moreover, such speech samples may be perceived as nonnatural, inasmuch as pauses of that size are rarely encountered in normal running speech. In searching for a preferential orientation pattern, we reasoned that it might be worthwhile to conduct an experiment with the same speech samples but with a systematic reduction in the length of the pauses. As the effect of pause length on infants' speech-perception abilities as tested by a preferential head-turn paradigm to our knowledge has not been investigated—pause lengths have often not even been reported (e.g., Jusczyk, Cutler, & Redanz, in press)—we decided to run an additional experiment to evaluate this issue.

EXPERIMENT 5

Experiment 5 tested the hypothesis that in the previous experiments pauses between the triplets were too long to keep the infant's attention continuously directed towards the speech signal. This experiment used the same speech samples as in Experiment 4, with the pauses reduced in length.

Method

Subjects. Twenty-four 9-month-old infants (mean age = 9.0 months; range 8.7 to 9.7) participated in this experiment. They all came from monolingual Dutch-speaking families, and they were healthy and had uncomplicated prenatal and postnatal histories. To obtain the experimental group, 36 subjects were tested. Twelve were excluded, 3 for crying, 4 for restlessness, 3 for too-short looking times, 1 for ignoring one side, and 1 because of computer failure.

Stimulus material. The materials were the same as those used in Experiment 4, except that the interstimulus interval between the triplets was reduced from 1,250 to 800 msec.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

Mean orientation times were 6.9 sec for the legal clusters and 6.1 sec for the illegal clusters. A t test did reveal a significant effect between the two conditions [$t(23) = 2.095$, $p < .05$]. Seventeen of the 24 infants preferred the legal structures.

Discussion

The results indicate that under very constrained context conditions infants at the age of 9 months are able to detect phonotactically marked word boundaries at the end of wordlike units, even when these are presented in connected speech. The particular conditions under which this

is possible for infants at that age can be inferred from the combined results of Experiments 2 through 5 concerning not only linguistic but also attentional parameters. We will discuss this below. Before doing so, however, we will briefly report the results from an additional experiment (Experiment 6) which, using the same interstimulus interval as Experiment 5, served as a replication of Experiment 1 in order to provide a more direct comparison with the infants' performance in Experiment 5.

EXPERIMENT 6

Experiment 6 served as a replication of Experiment 1 using the speech material of Experiment 1 but the interstimulus interval of Experiment 5.

Method

Subjects. Twenty-four 9-month-old infants (mean age = 9.0 months; range, 8.7 to 9.6) from monolingual Dutch-speaking families participated in this experiment; they were healthy and had uncomplicated prenatal and postnatal histories. To obtain the experimental group, 41 subjects were tested. Seventeen were excluded, 5 for crying, 5 for restlessness, 5 for too-short latencies, 1 because it was too tired, and 1 because of computer failure.

Stimulus material. The critical items from Experiment 1 were used in this experiment. In contrast to Experiment 1, however, the interstimulus interval between the isolated wordlike units was 800 msec. The total length of the speech trials in this experiment was 16 sec (mean lengths were 15 sec 380 msec for the legal speech samples and 5 sec 770 msec for the illegal speech samples). All other variables were the same as in Experiment 1.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

A main effect of condition was found [$t(23) = 2.43$, $p < .03$]. The mean looking times were longer for the legal structures (6.3 sec) than for the illegal structures (5.3 sec). Twenty of the 24 subjects showed a preferential pattern for the legal phonotactic structures.

Discussion

The results of the present experiment with an interstimulus interval of 800 msec between speech samples replicated the preferential pattern for legal over illegal phonotactic clusters in 9-month-old infants observed in Experiment 1, which had used an interstimulus interval of 1,250 msec. Infants showed a clear preference for the phonotactically legal over the phonotactically illegal items.

Thus, these data indicate that infants at the age of 9 months are sensitive to phonotactically marked word boundaries when wordlike units are presented in isolation regardless of the length of interstimulus interval between the critical units (Experiments 1 and 6). Moreover, the findings from Experiments 2 through 5 suggest that infants around the age of 9 months are able to use phonotactic knowledge about word boundaries to segment running speech, although only under a processing condition with reduced linguistic requirements (redundant context), reduced speech perception requirements (infant-directed speech mode), and reduced attentional requirements (reduced interstimulus interval).

One critical question that remained to be answered was whether the observed preferences in these experiments were indeed due to the infants' sensitivity to phonotactic cues of word boundaries or to some other experimentally uncontrolled suprasegmental cues. One way to test this is to use low-pass-filtered stimulus material, since low-pass filtering eliminates most of the phonetic and phonotactic information.

EXPERIMENT 7

In this experiment, we sought to determine if the observed preferences for legal over illegal word-boundary clusters were based on phonotactic cues or if there was sufficient information in the prosody of the spoken material to distinguish legal from illegal clusters. To do this, we low-pass-filtered the stimulus material at 400 Hz, thus eliminating most of the phonotactic information while preserving the prosodic features. It has been shown that when prosodic differences are salient, infants will continue to respond to them, even if the stimuli are low-pass filtered (Fernald & Kuhl, 1987; Jusczyk, Cutler, & Redanz, in press). Salient phonotactic differences, in contrast, should not be and are not detectable when the speech material is low-pass filtered at 400 Hz (Jusczyk, Friederici, et al., in press). For this reason, we decided to test 9-month-olds on a low-pass-filtered version of Experiment 6, expecting a nonpreferential pattern under the hypothesis that the observed preferences in the preceding experiments were due to phonotactic features.

Method

Subjects. Twenty-four 9-month-old infants (mean age = 9.2 months; range 8.7 to 9.4) from monolingual Dutch-speaking families participated in this experiment; they were healthy and had uncomplicated prenatal and postnatal histories. To complete the experimental group, 48 subjects were tested. Twenty-four were excluded, 4 for crying, 5 for restlessness, 8 for too-short latencies, 3 for lack of interest, 2 for missing items, and 2 for experimental failure.

Stimulus material. The stimulus material used in Experiment 6 was employed in the present experiment. However, the audio output was passed through a Krohn-Hite filter with its low-pass cutoff set to 400 Hz and an attenuation slope of 48 dB per octave.

Apparatus and Procedure. The apparatus and procedure were as in Experiment 1.

Results

With the low-pass-filtered stimulus material, no significant main effect of condition was found [$t(23) = 0.133$, $p = .895$]. The mean looking times were 6.5 sec for the legal clusters and 6.3 sec for the illegal clusters. Fifteen of the 24 subjects showed a preferential pattern for the legal phonotactic structures.

Discussion

The findings from this last experiment showed no preferential pattern for legal over illegal structures when

the stimulus material was low-pass filtered, strongly suggesting that those preferences that were observed in some of the previous experiments were due to the infants' sensitivity to the phonotactic cues of word boundaries and not to prosodic features.

GENERAL DISCUSSION

The present study has shown that during the second half of the first year of life, infants are acquiring knowledge about the phonotactic structure of words of their native language. The results of Experiment 1 and Experiment 6 clearly demonstrate that 9-month-old Dutch infants listen longer to speech samples whose word boundary structure is legal in Dutch than they do to speech samples whose word boundary structure contains clusters that are illegal in Dutch. The results of Experiment 5, moreover, indicate that 9-month-old infants are even capable of using this knowledge to identify word boundaries in a linguistically simple speech context. Nine-month-old infants orient longer to speech samples of three syllables whose middle syllable has a phonotactically legal word-offset structure than they do to those whose middle syllable contains phonotactically illegal word offsets. Remember that the illegality of the middle syllable was induced by the incorrect positioning of a legal word onset cluster at the end of a wordlike unit, resulting in a triplet the last two syllables of which could not be segmented in a phonotactically legal way. Thus, if infants are sensitive to this phonotactic manipulation, showing a preference for phonotactically legal over illegal triplets as in Experiment 5, we can infer that they must have segmented the speech triplets in a phonotactically adequate way. This, in turn, means that they are able to apply their knowledge about possible word boundaries to segment connected speech into wordlike units.

The results of Experiments 2 through 5 indicate that this sensitivity to legal over illegal word boundary structures in the context of connected speech can be observed only under certain circumstances in 9-month-olds. First, identification of word boundaries on the basis of phonotactic constraints seems possible only when the context structure is linguistically simple, for example, when it consists of two identical syllables. Second, word-boundary identification at that age appears to depend upon an infant-directed speech mode. These results are in agreement with earlier findings of Goodsitt et al. (1984), who reported that sensitivity to phonetic contrast in context can be observed in 6.5-month-old infants only when context syllables are redundant and when the speech mode is infant-directed. A comparison between Experiments 4 and 5, moreover, suggests that detection of the phonotactic violation of word boundaries in 9-month-old infants is not independent of attentional factors. A reduction in the pause length, from 1,250 to 800 msec, between triplets in one trial increased the infants' sensitivity to phonotactic variables in the speech samples. Other studies using language material consisting of speech trials containing word lists had not used pauses of constant length (e.g., Jusczyk, Cutler, & Redanz, in press; Jusczyk, Friederici, et al., in press). Variable pause lengths in these studies may have held the infants' attention better than monotonically reappearing pauses of constant length would have.

Experiment 7, using low-pass-filtered stimulus material, provides clear evidence for the claim that the observed preferences for legal over illegal phonotactic word-boundary clusters in the 9-month-olds are indeed due to phonotactic features and not to possible prosodic cues.

Related work by Jusczyk, Friederici, et al. (in press) has shown that infants around the age of 9 months, in contrast to 6-month-old infants, use phonotactic information to distinguish native-language from foreign-language words. The present findings show that around the same age infants seem to be sensitive to the phonotactic features that signal word boundaries in the native language. Moreover, the data suggest that infants at this age are able to use this knowledge to segment the speech stream into wordlike units. In particular, it was shown that phonotactically legal word onset clusters are perceived as illegal at the end of wordlike units when the critical syllable is embedded in a simple two-syllable context. This increase in sensitivity to phonotactic regularities in the native language during the first year of life goes together with a decrease in the sensitivity for those phonetic contrasts which are not part of the native language (Werker & Lalonde, 1988; Werker & Tees, 1984; but see also Best, McRoberts, & Sithole, 1988).

The speech-perception system gets tuned towards the characteristics of native language from birth (Mehler et al., 1988) and continues over the following years. Prosodic cues are used to identify clause and phrase boundaries between the ages of 6 and 9 months (Hirsh-Pasek et al., 1987; Jusczyk, 1989). At the age of 9 months, infants appear to be equipped with knowledge of possible word structures in their native language, be it at the prosodic level (Jusczyk, Cutler, & Redanz, in press) or at the phonotactic level (Jusczyk, Friederici, et al., in press). At this age, moreover, infants seem to be able to use phonotactic knowledge about possible word boundaries to identify wordlike units in connected speech, as shown in the present study.

Whether the observed sensitivity to phonotactic word-boundary constraints in 9-month-old infants can be taken to reflect fully developed phonotactic knowledge about possible words in their native language, or whether it must be characterized as a sensitivity to the frequency of occurrence of certain syllables in certain positions cannot be decided on the basis of the data at hand. It may well be that frequency of occurrence is the first ground upon which to build up initial language-specific knowledge, on the basis of which rule governed language knowledge will develop later.

In conclusion, the present results demonstrate that infants around the age of 9 months are able to detect phonotactically marked word boundaries in a linguistically simple context, and they suggest that infants around this age

are beginning to use knowledge about phonotactically possible word boundaries to segment the speech stream. The ability to identify word boundaries in context may develop fully only around the age of 11 months or later (Kemler Nelson, 1989; Werker & Pegg, 1992), and lexically bound phonemic knowledge may not be established before the age of 19 months (Werker & Pegg, 1992). The road towards a fully developed lexicon is long. The present data, however, suggest that a first milestone on the way towards a lexicon, that is, the ability to segment speech input into wordlike units, is present during the last 3 months of the first year of life.

REFERENCES

ASLIN, R. N. (1993). Segmentation of fluent speech into words: Learning models and the role of maternal input. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 305-315). Dordrecht: Klüwer.

ASLIN, R. N., PISONI, D. B., HENNESSY, B. L., & PEREY, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development*, **52**, 1135-1145.

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

CELEX (1990). *CELEX—A guide for users*. Nijmegen, The Netherlands: University of Nijmegen.

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.

EIMAS, P. D., MILLER, J. L., & JUSCZYK, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception* (pp. 161-195). New York: Cambridge University Press.

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

FERNALD, A., & KUHL, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development*, **10**, 279-293.

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.

GOODSITT, J. V., MORSE, P. A., VER HOEVE, J. N., & COWAN, N. (1984). Infant speech recognition in multi syllabic context. *Child Development*, **55**, 903-910.

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.

JUSCZYK, P. W., CUTLER, A., & REDANZ, N. J. (in press). Infants' sensitivity to the predominant stress pattern of English words. *Child Development*.

JUSCZYK, P. W., FRIEDERICI, A. D., WESSELS, J., SVENKERUD, V., & JUSCZYK, A. M. (in press). Infants' sensitivity to the sound structure of native language words. *Journal of Memory & Language*.

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.

KARZON, R. G. (1985). Discrimination of polysyllabic sequences by one-to-four-month-old infants. *Journal of Experimental Child Psychology*, **36**, 326-342.

KEMLER NELSON, D. G. (1989, April). *Developmental trends in infants' sensitivity to prosodic cues correlated with linguistic units*. Paper presented at the Biennial meeting of the Society for Research in Child Development, Kansas City.

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.

KUHL, P. K. (1987). Perception of speech and sound in early infancy. In P. Salapatek & L. Cohen (Eds.), *Handbook of infant perception* (pp. 275-381). New York: Academic Press.

KUHL, P. K., WILLIAMS, K. A., LACERDA, F., STEVENS, K. N., & LINDBLOM, B. (1992). Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science*, **255**, 606-608.

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 (Eds.), *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.

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 & Development*.

WERKER, J. F., & LALONDE, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, **24**, 1-12.

WERKER, J. F., & PEGG, J. E. (1992). Infant speech perception and phonological acquisition. In C. Ferguson, L. Menn, & C. Steol-Gammon (Eds.), *Phonological development: Models, research and implications* (pp. 285-311). Parkton, MD: York Press.

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

(Manuscript received August 21, 1992;
revision accepted for publication February 18, 1993.)