

# Infant-Directed Speech Facilitates Word Segmentation

Erik D. Thiessen  
*Department of Psychology*  
*Carnegie Mellon University*

Emily A. Hill and Jenny R. Saffran  
*Department of Psychology and Waisman Center*  
*University of Wisconsin–Madison*

There are reasons to believe that infant-directed (ID) speech may make language acquisition easier for infants. However, the effects of ID speech on infants' learning remain poorly understood. The experiments reported here assess whether ID speech facilitates word segmentation from fluent speech. One group of infants heard a set of nonsense sentences spoken with intonation contours characteristic of adult-directed (AD) speech, and the other group heard the same sentences spoken with intonation contours characteristic of ID speech. In both cases, the only cue to word boundaries was the statistical structure of the speech. Infants were able to distinguish words from syllable sequences spanning word boundaries after exposure to ID speech but not after hearing AD speech. These results suggest that ID speech facilitates word segmentation and may be useful for other aspects of language acquisition as well. Issues of direction of preference in preferential listening paradigms are also considered.

In the first years of life, infants must discover the patterns and structures that govern their native language. In the absence of a developmental disorder, infants regularly succeed at this task, eventually producing and comprehending language in an adult manner. Although infants eventually learn a language that is indistinguishable from that of other members of their linguistic community, they are exposed, during their formative years, to examples of language that are often quite different. This is due to the fact that adults speak to infants differently than they speak to

other adults. Overall, infant-directed (ID) speech, in comparison to adult-directed (AD) speech, is characterized by a slower rate of speech, a higher fundamental frequency, greater pitch variation, longer pauses, characteristic repetitive intonational structures, and simplified sentence structure (e.g., Fernald, 1992; Gleitman, Newport, & Gleitman, 1984).

Even though ID speech is in some ways unlike the language that infants will speak as adults, the ways in which ID speech differs from AD speech may actually facilitate infants' learning. For example, novel or focused words are frequently placed at the ends of utterances in ID speech (Fernald & Mazzie, 1991). Infants are most successful at recognizing (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998) and segmenting (Aslin, 2000) utterance-final words. Thus, ID speech may facilitate lexical comprehension and word learning. Further, the simplified phrasal structure and exaggerated prosodic marking of phrases in ID speech provide redundant cues to grammatical structure, and such redundancy may aid learning (e.g., Fisher & Tokura, 1996; Morgan, Meier, & Newport, 1987; Steedman, 1996; Venditti, Jun, & Beckman, 1996). Consistent with this hypothesis, Fernald and Cummings (2003) found that simple repetitive word sequences—an experimental analog of ID speech's simplified phrase structure—facilitated infants' word recognition. Finally, ID speech may help to focus infants' attention on language, which could speed learning (e.g., Rose, Feldman, & Jankowski, 2003).

Although ID speech captures infants' attention more readily than AD speech (e.g., Cooper & Aslin, 1990; Fernald, 1985), there is little direct evidence to suggest that ID speech facilitates language learning. Indeed, children's lexical access does not appear to be faster when a word is presented in ID speech than in AD speech (Fernald & Cummings, 2003). However, Golinkoff and Alioto (1995) found that English-speaking adults learned novel object names more easily when those names were presented in ID speech rather than in AD speech. This result is consistent with the suggestion that ID speech may assist language learners in segmenting the speech stream or in forming links between words and referents. However, adults are a very different population than infants. The fact that adults learn aspects of a novel language more easily when exposed to ID speech than AD speech does not necessarily mean that ID speech will also facilitate infants' learning. To assess this possibility, we must ask whether ID speech helps infants learn about some aspect of language. The task Golinkoff and Alioto used—word learning—is not appropriate, because young infants do not learn the meanings of words very easily. However, young infants do learn to identify words in fluent speech (e.g., Jusczyk & Aslin, 1995). Therefore, it is possible to explore whether infants segment words more easily from fluent ID speech than from fluent AD speech.

Because word boundaries in speech are not consistently marked by pauses (unlike the white spaces in text), infants must use other aspects of the speech stream to discover words in fluent speech. One such cue is the statistical structure of speech: Syllable combinations that are part of the same word tend to occur together more

consistently than syllable combinations spanning word boundaries (Hayes & Clark, 1970). As an example, consider the phrase *pretty baby*. Whereas the syllable *pre* is highly predictive of the syllable *ty* (because there are few words in English that begin with *pre* other than *pretty*), *ty* does not predict *ba* very strongly, as there are many other syllables that can follow *ty* (e.g., *pretty eyes*, *pretty hair*, *pretty dress*). Therefore, the transitional probability from *pre* to *ty* is high, whereas the transitional probability from *ty* to *ba* is low. Infants can distinguish between syllable sequences with high and low transitional probabilities and readily use this type of statistical information to discover word boundaries (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996). The experiments reported here are designed to discover whether infants' use of this information is facilitated by ID speech.

There are (at least) two possible ways in which ID speech may facilitate learning. First, ID speech may provide infants with more information, as previous acoustic and linguistic analyses of this register have indicated (e.g., Fernald & Mazzie, 1991; Kuhl et al., 1997). For example, the longer pauses at phrase boundaries in ID speech could make it easier for infants to locate word boundaries. Second, ID speech may enable infants to more efficiently discover or use information in speech, even in situations when it conveys the same information (e.g., statistical information) as AD speech. That is, if infants were exposed to two languages—each with the same statistical information—would they learn more easily from ID speech than from AD speech?

It is this second possibility we explore in this series of experiments. Our question is not whether the acoustic characteristics of ID speech are such that word boundaries are more pronounced in ID speech than in AD speech. Rather, our question is whether the characteristics of ID speech aid use of statistical cues. To answer this question, it is impossible to use completely naturalistic ID and AD speech, because ID speech may contain more auditory cues to word boundaries (i.e., it makes the word segmentation problem easier) than AD speech. Instead, we created ID speech that included only one characteristic of natural ID speech: its prosodic characteristics. Importantly, the exaggerated pitch contours in the ID speech do not indicate word boundaries; the only cue to word boundaries in the speech stream is its statistical structure (which is identical in the ID and AD speech).

Although instances in which both ID and AD speech convey the same information may be rare in infants' natural environments, by equating them in this manner, we can better understand whether ID speech is facilitative (if it is) because it provides more information, or whether it allows for better learning even when both ID and AD speech provide the same information. By comparing infants' success at segmenting the ID speech to their performance segmenting the same words from AD speech, we can determine whether the enhanced prosodic characteristics of ID speech facilitate the detection of statistical cues to word boundaries. If we find that

infants more successfully use statistical cues in ID speech than in AD speech—even when both types of speech contain the same statistical cues—then we may learn more than simply whether or not ID speech facilitates learning. We may also begin to understand why ID speech is facilitative.

## EXPERIMENT 1

We created two artificial languages in this experiment. Each language consisted of 12 sentences, made up of nonsense words. One set of sentences was spoken in an AD speech register, whereas the other set was spoken with the exaggerated prosody of ID speech. There are a variety of ways in which the prosodic characteristics of ID speech differ from those of AD speech. Katz, Cohn, and Moore (1996) argued that both dynamic pitch contour and summary measures of pitch (e.g., mean and standard deviation of fundamental frequency) are necessary to completely describe ID speech. For example, ID speech has very characteristic pitch contours (e.g., Fernald, 1989). These different contours also have different average fundamental frequency (F0). The different aspects of the prosodic characteristics of ID speech can have different effects. Trainor and Desjardins (2002) suggested that although the exaggerated pitch contours of ID speech should improve infants' abilities to discriminate between vowels, the higher average F0 of ID speech would not help infants make such distinctions and might actually impair their ability to make them. In this experiment, the only difference between ID speech and AD speech was its prosodic characteristics. However, there is no way to be certain which of these characteristics—if any—will have an effect on learning and what that effect might be. As such, in this experiment we attempted to mimic the prosodic characteristics of natural ID speech as faithfully as possible. If we find that our ID speech facilitates learning, it will remain an open question which aspect of our ID speech's pitch characteristics (e.g., exaggerated contour, higher F0) is responsible for such facilitation.

Although the ID speech used in this experiment is not fully natural—because our ID speech has only the prosodic characteristics of natural ID speech, and none of the other auditory or structural characteristics that differentiate ID speech from AD speech—these results are relevant to understanding possible benefits of natural ID speech. This series of experiments will allow us to assess whether the prosody of ID speech facilitates learning by enabling infants to more efficiently discover or use the statistical structure of ID speech than AD speech. If infants in this experiment learn more easily from ID speech than from AD speech, then infants should also learn more easily from natural ID speech, because natural ID speech has the same prosodic characteristics as the speech used in this experiment.

To assess whether infants learn from ID speech more easily than AD speech, we familiarized two groups of infants with sentences containing novel nonsense

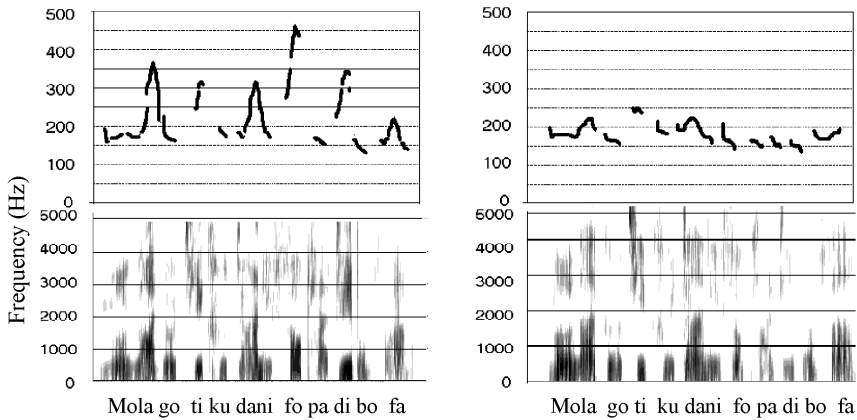
words. One group heard the sentences spoken with a pitch range and intonational structure characteristic of ID speech, whereas the other group heard the same sentences spoken in a more monotonic fashion characteristic of AD speech. In neither case did prosodic cues mark word boundaries; the only cues to word boundaries in both conditions were the statistical properties of speech. After exposure to these sentences, we tested whether infants could discriminate between syllable sequences that were words in the languages they heard and syllable sequences that were part words (syllable sequences that crossed word boundaries; e.g., *tyba* in the phrase *pretty baby*). The question of interest was whether infants who heard the ID speech performed the discrimination between words and part words more easily than infants who heard the AD speech.

## Method

*Participants.* Participants were 40 infants between the ages of 6.5 and 7.5 months. Twenty infants participated in the ID speech condition ( $M = 7.07$  months), and 20 participated in the AD speech condition ( $M = 7.12$  months). To obtain these 40 infants, it was necessary to test 54. The other 14 (7 from each condition) were excluded for the following reasons: fussiness (7), average looking time of less than 3 sec to one or both side lights (3), failure to complete at least 8 (of 12) test trials (3), and parental interference (1). According to parental report, all infants were full-term and free of ear infections at the time of testing.

*Stimuli.* The familiarization language consisted of 12 different nonsense sentences, each composed of four nonsense words (*dibo, kuda, lagoti, nifopa*). Each sentence began with the buffer syllable *mo*, and ended with the buffer syllable *fa*, to ensure that participants could not use the silent pause before and after each sentence as a cue to word boundaries (see the Appendix for a list of all 12 sentences). The only cues to word boundaries were the probabilities with which syllables co-occurred (1.0 within words vs. .25 at word boundaries). The sentences were spoken by a female native English speaker naive to the placement of word boundaries; her script consisted only of strings of 12 syllables with no word boundary markings. All sentences were completely coarticulated and spoken fluently. There were no extraneous pauses between words or syllables in the final language.

The same speaker produced both sets of sentences (ID and AD). She produced several tokens of each of the 12 nonsense sentences, each with different intonation contours and different syllables stressed. After the initial recordings, the ID and AD speech differed not only with respect to their prosodic characteristics (e.g., pitch contour, F0) but their rate and amplitude as well (because it is difficult for a speaker to adjust pitch while leaving rate and amplitude unaffected). As such, the ID speech was initially both longer and louder than the AD speech. To remove these possible confounds, the length and amplitude of both sets of sentences were



**FIGURE 1** F0 contour and spectrogram of the sentence “mo lagoti kuda nifopa dibo fa” in infant-directed (left) and adult-directed (right) speech. Both sentences have durations of 4.2 sec.

adjusted to an average median value, using CoolEdit’s editing tools. After this editing, both the ID and AD speech had the same overall amplitude (approximately 60 dB at the infant’s head) and rate (2.5 syllables/sec). Both languages consisted of 12 sentences, with pauses of 1.3 sec between sentences. Each set of 12 sentences was 1 min in duration. Representative examples of one ID sentence and one AD sentence are shown in Figure 1.

In the AD speech, the speaker’s average F0 was 230 Hz (range = 140–260 Hz). In the ID speech, the speaker’s average F0 was 292 Hz (range = 140–480 Hz). The increased range in the ID speech was due to the speaker’s exaggerated pitch peaks, which reached an average of 406 Hz. By contrast, pitch peaks in the AD speech reached an average of only 252 Hz. These values are consistent with previous reports on the F0 characteristics of ID speech. For example, Fernald (1989) found that the average F0 of ID speech was 311 Hz, with a range of 338 Hz.

Pitch peaks are a prominent cue to stress (e.g., Sluijter, van Heuven, & Pacily, 1996), which is a strong indicator of word boundaries in English. That is, in English, pitch peaks are more likely to occur on word-initial syllables. To ensure that pitch peaks did not mark word boundaries in these materials (especially in the ID speech), we equalized the number of times each syllable within a word occurred on a pitch peak in both the ID and AD speech. To do so, we selected sentences—from among the numerous tokens of each sentence our speaker recorded—that ensured that no syllables would be consistently stressed across the 12 sentences. After the selection process, pitch peaks were distributed evenly across the syllables in words, as judged by two listeners and by visual inspection of the F0 contour. This process ensured that pitch peaks did not occur more frequently on word-initial syllables than on other positions within a word.

Four test items were recorded in isolation, two of which were words from the language (*dibo* and *lagoti*) and two of which were part words, syllable sequences that crossed word boundaries (*paku* and *danifo*). All test items were spoken in a monotonic fashion similar to the AD speech, with a mean F0 of 208 Hz.

For this experiment to be a valid comparison of infants' learning from ID and AD speech, it is important that both types of speech be equally naturalistic and that our ID speech sounds as though it is ID. To address these points, we asked 10 adults to listen to the stimuli in these experiments. Each adult heard all 12 ID sentences in a block and all 12 AD sentences in a block. Half of the adults heard the block of ID speech first, and half of the adults heard the block of AD speech first. After listening to each group of sentences, adults were asked to judge how natural the sentences sounded on a scale of 1 (*very unnatural*) to 7 (*very natural*), and how likely each group of sentences were to be spoken to children, also on a scale from 1 (*very unlikely to be spoken to children*) to 7 (*very likely*).

After listening to the ID speech, adults rated its naturalness as 4.6 ( $SE = 0.3$ ). After listening to the AD speech, adults rated its naturalness as 5.0 ( $SE = 0.1$ ). This difference was not significant,  $t(9) = 0.89$ ,  $p > .05$ . Although adults rated both kinds of sentences as equally natural, there was a sizable difference in how likely adults thought they were to be spoken to children. Adults, on average, gave the ID speech a rating of 6.2 ( $SE = 0.2$ ) on this measure, whereas AD speech was given a rating of only 3.5 ( $SE = 0.4$ ). This difference was significant,  $t(9) = 6.96$ ,  $p < .01$ . Thus, although both types of speech were judged to be equally natural, adults judged the ID speech to be more likely to be addressed to children. This suggests that although our stimuli do not match every characteristic of ID speech (e.g., simplified phrase structure), the prosodic characteristics of our ID speech allow it to approximate natural ID speech.

Because the sentences in this experiment were spoken by a natural speaker—instead of a synthesizer—it is possible that additional cues to word boundaries may have unintentionally been introduced into the ID speech. Although the acoustic analyses just reported suggest that some of the most obvious cues were not present, it is nevertheless possible that more subtle word boundary cues were differentially available in the two conditions. To determine whether auditory cues (i.e., cues other than sequential statistics) in the ID speech indicated word boundaries, 20 native adult participants (who had not previously assessed the stimuli for naturalness or likelihood of being addressed to a child) were asked to listen to one sentence—of six possible sentences in which the part words that were used as test items occurred—from our ID speech, repeated as often as they needed. Participants were then asked to identify which of two items—a word and a part word—was more likely to have been a word in the sentence they heard. Each participant heard a sentence selected at random. Our rationale was as follows: If there were acoustic cues to word boundaries in the sentences, adults should be able to detect them by listening (especially listening an unlimited number of times). However, if only statistical



cues indicated word boundaries in the sentences, adults should not be able to distinguish between words and part words, because given just one sentence, all of the transitions between syllables (both within a word and across word boundaries) occur with 100% probability.

After listening to the sentence, adults were presented with a test trial in which they heard a word and a part word (the same test items presented to the infants). Participants were asked to identify which item—the first or the second—had been a word in the sentences they heard. Adults were correct 45% of the time, a rate that is not significantly different from chance (50%),  $t(19) = 0.44$ ,  $p > .05$ . These results indicate that the ID speech indeed lacked acoustic cues to word boundaries in our speech, such that only sequential statistical cues were available for word segmentation.

*Procedure.* Infants were tested individually in a double-walled, sound-attenuated room while seated on a parent's lap. An experimenter outside the booth observed the infants' looking behavior on a video monitor connected to a camera inside the room, and coded the direction of the infants' gaze in real time. To eliminate bias, the parent inside the room listened to masking music over headphones, and the experimenter was similarly unable to hear the stimulus being played to the infant.

At the beginning of the familiarization phase, a light in the center wall facing the infant began to flash, directing the infant's gaze forward. Simultaneously, one of the two sentence sets (either AD or ID; each infant heard only one set) began to play from the speakers beneath the two side lights—one light and speaker on each side wall—in the room. The familiarization phase lasted for 1 min.

Immediately after familiarization, 12 test trials were presented. All infants heard the same test trials, regardless of familiarization condition. Six of the trials were word trials, and 6 of the trials were part-word trials. Test items were presented in random order, with 6 trials (3 word trials and 3 part-word trials) presented from each side speaker. Each test item occurred on 3 trials during the testing session. A test trial began with a flashing light at the center of the wall facing the infant, drawing the infant's gaze forward. When the experimenter signaled the computer that the infant had fixated on the center light, one of the side lights began to flash, and the center light simultaneously stopped. As soon as the infant made a head turn of at least 30° in the direction of the flashing side light, the experimenter signaled the computer, and one of the test items was presented from the speaker beneath the flashing side light. When the infant looked away for more than 2 sec, the test item stopped playing, and the center light began to flash again. This procedure continued for as long as the infant was willing to attend or until they had completed all 12 test trials.

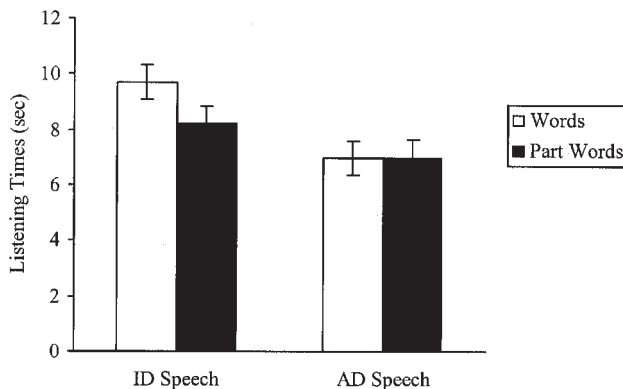


## Results and Discussion

First, we compared listening times to words and part words for infants exposed to the ID speech. As shown in Figure 2, infants listened to words for 9.7 sec ( $SE = 0.7$ ) and to part words for 8.2 sec ( $SE = 0.7$ ). Fourteen of the 20 infants listened longer to words than part words during the test trials. A paired  $t$  test (all  $t$  tests reported are two-tailed) indicated that the difference in looking times between words and part words was significant,  $t(19) = 2.09, p < .05$ .

Second, we compared listening times to words and part words for infants exposed to the AD speech. As shown in Figure 2, infants listened to words for 7.0 sec ( $SE = 0.5$ ) and to part words for 7.0 sec ( $SE = 0.6$ ). Nine of the 20 infants listened longer to words than part words during the test trials. A paired  $t$  test indicated that the difference between infants' listening times to words and part words was not significant,  $t(19) = 0.01, p > .05$ .

In the ID speech condition, infants showed a preference for words, whereas in the AD speech condition, infants listened equally to words and part words. To assess whether this was a significant difference, we performed a  $2 \times 2$  (Condition  $\times$  Item) analysis of variance (ANOVA). There was a main effect of condition (ID vs. AD),  $F(1, 38) = 5.67, p < .05$ , because infants' looking times were longer overall after exposure to the ID speech. This may have been due to the fact that the test items were spoken in a monotone similar to AD speech, making them dissimilar to the ID speech and thus generally more interesting to infants in the ID speech condition. There was a nonsignificant trend toward a main effect for item (word vs. part word),  $F(1, 38) = 3.35, p = .08$ . Most important, there was also a trend toward an interaction,  $F(1, 38) = 3.35, p = .08$ , which suggests that infants' preference for



**FIGURE 2** Seven-month-old infants' listening times to words and part words after exposure to infant-directed and adult-directed speech in Experiment 1.

words over part words after exposure to the ID speech was different from their lack of preference after exposure to the AD speech.

These results are consistent with the hypothesis that infants found the ID speech easier to segment than the AD speech. After exposure to the ID speech, infants discriminated between words and part words. To be able to discriminate words from part words, infants must have learned something about words by segmenting the fluent speech via statistical cues. Infants showed no evidence that they distinguished words from part words after hearing the AD speech. This suggests that infants did not segment the AD speech and thus were unable to detect the difference between words and part words.

However, there is another explanation for these results based on what is known about the factors that predict infant direction of preference (e.g., Hunter & Ames, 1988). In general, direction of preference lies on a continuum from familiarity preferences to novelty preferences (e.g., Wagner & Sakovits, 1986). An infant's position on the continuum depends on the difficulty of the task. If the task is difficult, infants will show a familiarity preference and listen longer to items that match or are reminiscent of the familiarization stimuli. As the task gets easier (e.g., due to length of familiarization, stimulus complexity, age, etc.), infants move toward a novelty preference due to habituation. If the task is relatively easy, infants will listen longer to items that are less reminiscent of the familiarization stimulus. On tasks of intermediate difficulty, infants are balanced between familiarity and novelty and may show no consistent direction of preference.

One of the factors that predict task difficulty is the degree of similarity between familiarization stimuli and test stimuli (Hunter & Ames, 1988; Thiessen & Saffran, 2003). If the test stimuli closely match the familiarization stimuli, the task is easier for infants than if there is only a distant match between familiarization and test items. In this experiment, the test items were spoken in a relatively monotonic fashion. That is, the test items were more similar to the AD speech than to the ID speech.

Therefore, another possible explanation for the results of Experiment 1 is that infants found both the ID speech and the AD speech equally easy to segment. However, the infants who were exposed to the ID speech found the test trials difficult because the test items did not closely match the familiarization materials with respect to their pitch properties. On this view, infants in the ID speech condition showed a familiarity preference due not to ease of learning but to the difficulty of the test items. By contrast, infants in the AD speech condition may have found the test trials easy because the prosodic characteristics of the test trials closely matched the AD familiarization language. Because of this, infants in the AD speech condition may have been moving toward a novelty preference. It is thus possible that infants in the AD speech condition showed no preference because they were balanced between a familiarity preference and a novelty preference. If this is the case, then there is no evidence that ID speech facilitated learning.

As these considerations demonstrate, it is not possible to draw strong conclusions about the relative difficulty of two learning conditions based solely on a comparison between a familiarity preference in one condition and no preference in the other. Such a pattern of results could indicate that infants are learning in the condition that results in a familiarity preference, whereas infants in the other condition are not learning (and thus show no preference). However, this pattern could also indicate that infants are learning in both conditions and that the condition that results in no preference is the easier of the two. If this were the case, the lack of preference would indicate that infants are in the process of switching from a familiarity preference to a novelty preference.

Because the infants who heard the ID speech showed a familiarity preference, the current data do not allow us to assess the validity of the hypothesis that infants segment words more easily from ID speech than from AD speech. However, if infants in the ID speech condition were to show a novelty preference instead, it would be possible to draw conclusions about the relative difficulty of the AD speech condition. If infants in both conditions showed a novelty preference, we could conclude that the AD speech is as easy to segment as the ID speech (or at least, that infants did not demonstrate a difference in their degree of difficulty). However, if infants in the AD speech condition show either no preference or a familiarity preference, whereas infants in the ID speech condition show a novelty preference, our results would indicate that the AD speech was more difficult to segment. We test these possibilities in the next experiment.

## EXPERIMENT 2

Experiment 2 is a conceptual replication of Experiment 1, designed to elicit a novelty preference to facilitate cross-condition comparisons. To do so, we made two changes to the procedure of Experiment 1 intended to make the segmentation task easier, increasing the likelihood of a novelty preference. First, we doubled the length of the familiarization. Second, we recruited older participants. Both of these changes should ensure that infants in this experiment find the segmentation task easier than did the infants in Experiment 1.

### Method

*Participants.* Participants were 50 infants between the ages of 7.5 and 8.5 months. Of these, 25 participated in the ID speech condition ( $M = 7.97$  months), and 25 participated in the AD speech condition ( $M = 7.92$  months). To obtain these 50 infants, it was necessary to test 63. The other 13 were excluded for the following reasons: fussiness (6), average looking times of less than 3 sec to one or both side lights (4), failure to complete at least two of one or more trial types (2), and paren-

tal interference (1). Seven infants were excluded from the ID speech condition, and 6 infants were excluded from the AD speech condition.

*Stimuli.* The stimuli were identical to those used in Experiment 1. The only difference was that infants heard the set of 12 sentences twice, rather than only once, during familiarization. Each set of 24 sentences was 2 min in duration.

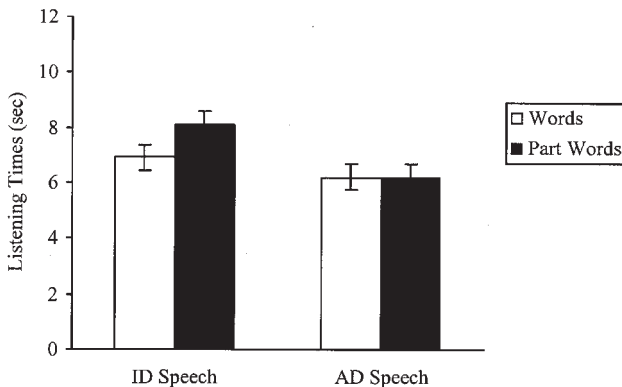
*Procedure.* The procedure was identical to that used in Experiment 1.

## Results and Discussion

As shown in Figure 3, after exposure to the ID speech, infants listened to words for 6.9 sec ( $SE = 0.6$ ) and to part words for 8.1 sec ( $SE = 0.5$ ). Nineteen of the 25 infants listened longer to part words than to words during the test trials. A paired  $t$  test indicated that the difference in looking times between words and part words was significant,  $t(24) = 2.51, p < .05$ .

After exposure to the AD speech, infants listened to words for 6.2 sec ( $SE = 0.4$ ) and to part words for 6.2 sec ( $SE = 0.5$ ). Twelve of the 25 infants listened longer to part words than to words. A paired  $t$  test indicated that the difference in looking times between words and part words was not significant,  $t(24) = 0.11, p > .05$ .

Infants who heard the ID speech distinguished between words and part words, whereas infants who heard the AD speech did not. Further, infants in the ID speech condition showed a novelty preference. It is impossible to draw conclusions about the relative difficulty of two conditions when infants in one condition show a familiarity preference and infants in the other show no preference, as in Experiment 1. However, a novelty preference indicates an easier task than results indicating ei-



**FIGURE 3** Eight-month-old infants' listening times to words and part words after exposure to infant-directed and adult-directed speech in Experiment 2.

ther no preference or a familiarity preference (Hunter & Ames, 1988). Therefore, the fact that infants show a novelty preference after hearing ID speech, but no preference after hearing AD speech, suggests that they found the ID speech easier to segment.

To determine whether infants' performance was significantly different across conditions, we performed a  $2 \times 2$  (Condition  $\times$  Item) ANOVA. As in Experiment 1, there was a main effect of condition (ID vs. AD),  $F(1, 48) = 4.84, p < .05$ , because infants' looking times were longer overall after exposure to the ID speech. Again, this may have been due to the fact that the test items were spoken in a monotone, making them dissimilar to the ID speech and thus more interesting to infants in the ID speech condition. There was a nonsignificant trend toward a main effect for item (word vs. part word),  $F(1, 48) = 3.64, p = .06$ . Most important, there was a significant interaction,  $F(1, 48) = 4.18, p < .05$ , showing that infants' preference for part words over words after exposure to the ID speech was significantly different from their lack of preference after exposure to the AD speech. These results support the hypothesis that ID speech is easier to segment than AD speech.

## GENERAL DISCUSSION

Our results are the first to show that the prosody of ID speech facilitates some aspect of language acquisition—in this case, word segmentation. Whereas infants were able to discriminate between words and part words after exposure to the ID speech in both experiments, they showed no evidence of this distinction after exposure to the AD speech. These results indicate that infants successfully segmented the ID speech but failed to segment the AD speech. Importantly, the ID speech did not provide additional cues to word boundaries beyond the statistical cues available in both the ID and AD speech. The prosodic characteristics of the ID speech apparently assisted in the acquisition of purely statistical information: the probabilities with which syllables co-occurred in sequence.

This is a different process than those that have been previously hypothesized to explain possible facilitative effects of ID speech (e.g., Gleitman et al., 1984). A number of examinations of ID speech have suggested that it may actually be more informative than AD speech. That is, the structure of ID speech may make linguistic patterns easier to discover than AD speech. For example, the simple sentence structure of ID speech may make it easier for infants to learn about certain aspects of word order. However, in this experiment, the structure of the ID and AD speech was identical. Each language had the same sentence structure and word order. Instead of containing more information, the ID speech made it easier for infants to access information also available in the AD speech. These results should not be construed to mean that ID speech is not, in the infants' natural environment, more

informative. Instead, they suggest an additional manner in which ID speech may be facilitative.

These results also do not imply that infants are unable to learn from AD speech; several prior experiments have demonstrated that infants can successfully segment entirely monotonic speech streams. In particular, Saffran et al. (1996) demonstrated that infants could segment monotonic synthesized speech based only on statistical cues. In Saffran et al.'s experiment, though, infants heard each word 45 times. In this experiment, infants heard each word only 12 times (in Experiment 1) or 24 times (in Experiment 2). Based on Saffran et al.'s results, and other experiments involving similar amounts of exposure to monotonic speech (e.g., Aslin et al., 1998; Johnson & Jusczyk, 2001), we expect that infants would have successfully segmented the AD speech with more exposure. As such, these results do not suggest that infants are unable to segment AD speech; instead, they indicate that ID speech may facilitate language acquisition by enabling faster or more efficient learning. The facilitative effect ID speech has on word segmentation may be slight. Indeed, in cultures in which adults do not employ ID speech, infants acquire language perfectly well (e.g., Schiefellin, 1990)—although there may be other characteristics of the language environment in these cultures that provide a facilitative effect similar to that of ID speech. Although the facilitative effect of ID speech may be slight, even a small increase in the speed or efficiency of initial acquisition can have large effects on the final characteristics of a learned system (e.g., Elman et al., 1996).

Word segmentation is only one aspect of language acquisition, and it is quite possible that ID speech facilitates other aspects of acquisition (e.g., Gleitman et al., 1984; Liu, Kuhl, & Tsao, 2003). Moreover, these results likely underestimate the facilitative effects of ID speech because our ID speech was equated to the AD speech on many of the dimensions that may make ID speech facilitative. Both the ID and AD speech had the same rate, pause duration, and sentence structure. Importantly, the effectiveness of our pitch manipulation is consistent with previous research that indicates that infants' preference for ID speech over AD speech is primarily due to the pitch characteristics of ID speech, rather than its amplitude or duration (Fernald & Kuhl, 1987). However, it is not clear which aspect of the prosodic characteristics of our ID speech—for example, average F<sub>0</sub>, F<sub>0</sub> range, or F<sub>0</sub> contour—or several of these characteristics in concert were responsible for infants' better learning in the ID speech condition.

Although our results indicate that some aspect of the prosodic characteristics of ID speech facilitate infants' access to the statistical structure of speech, they do not specify the exact mechanism through which this occurs. One possibility is that ID speech attracts and sustains infants' attention better than AD speech. There is a large body of research consistent with the claim that ID speech is more likely to hold infants' attention than AD speech (e.g., Werker, Pegg, &

McLeod, 1994). Recent research indicates that this attention-grabbing phenomenon is not limited to ID speech. Infants' preference for ID speech is similar to their preference for positive highly emotional AD speech (e.g., Singh, Morgan, & Best, 2002). Trainor, Austin, and Desjardins (2000) speculated that ID speech appears acoustically different from AD speech primarily because speech directed toward infants displays much more emotion than speech typically directed toward adults. Although it is not clear which aspect of ID speech attracts infants' attention, it is clear that it does so. This could, in turn, affect learning, as attention has been found to aid learning in a variety of tasks. For example, Frick and Richards (2001) found that infants were better able to later recognize a stimulus if the stimulus was presented to them while they were in a phase of sustained attention (as defined by heart rate).

In this experiment, there are several ways in which more sustained attention could promote learning. The ID speech could make infants more interested in the experimental situation in general, and thus lead to them performing more reliably during test trials. That is, infants may have learned equally well in both the ID and AD speech condition, but because their interest in the experimental room itself was better sustained in the ID speech condition, infants showed their learning more reliably in the AD speech condition. On this account, infants became bored in the AD speech condition, and thus did not attend to the test trials (even though they had learned the words). If this were the case, though, we would expect more infants to fuss or cry in the AD speech condition, a pattern we did not see in the data. An alternative hypothesis is that infants' increased attention to the ID speech may have improved their memory for the items they segmented from fluent speech, which allowed them to more easily discriminate between words and part words. If this account were correct, it would suggest that infants in both conditions segmented the fluent speech equally well, but only the infants who heard the ID speech remembered what they had segmented. This hypothesis is consistent with a body of research suggesting that attention can improve memory (e.g., Hertel & Rude, 1991; Rensink, O'Regan, & Clark, 2000).

There is a third possible way in which increased attention may have affected infants' performance in this task. Instead of keeping infants from boredom with the experimental setting, the ID speech may have increased infants' attention to the fluent speech itself, and thus to the statistical relations between syllables. Ideally, to discover the statistical structure of a speech stream, a listener would attend to—and remember—every single instance of each syllable and its co-occurrences with the syllables preceding and following it. In general, the further a listener is from this ideal state, the worse he or she will be at detecting the statistical structure; in the extreme case where the listener fails to attend to any of the syllables, he or she will completely fail to discover the statistical relations between syllables. Perhaps infants who heard ID speech were more easily able to discover relations between



syllables (and thus to distinguish between words and part words) because they were more attentive to the speech. On this account, infants in the ID speech condition segmented the fluent speech successfully, whereas infants who heard the AD speech failed to do so.

Kaplan and his colleagues, in their research on infants' associative learning in response to ID and AD speech, suggested a similar hypothesis. Kaplan, Bachorowski, and Smoski (2002) found that infants of chronically depressed mothers learned pairings between faces and speech more easily when the speech consisted of typical ID speech than when the speech consisted of samples of depressed mothers' ID speech, which is less prosodically variable than typical ID speech (Kaplan, Bachorowski, & Smoski, 2001). In addition, Kaplan, Jung, Ryther, and Zarlengo-Strouse (1996) found that 4-month-old infants were more successful learning an association between a visual stimulus (a face) and ID speech than an association between a face and AD speech. Kaplan et al. (1996) suggested that ID speech may sensitize infants to the environment around them and prime them to notice relations between events, which enables infants to learn associations more easily. Although Kaplan et al.'s (1996) paradigm is very different, their hypothesis is relevant to the experiments reported here because using transitional probabilities to segment words from fluent speech involves detecting which syllables are associated (occur together in words) and which are less strongly associated (occur together only at word boundaries).

Each of these hypotheses is consistent with the current data. Therefore, these experiments do not distinguish between them. However, they do indicate that ID speech enables infants to more easily use statistical information than AD speech, even in situations where AD and ID speech contain the same statistical structure. Thus, these results indicate that ID speech may do more than provide infants with information that AD speech does not. ID speech may also allow infants to more easily learn from information present in speech. As such, these results provide a compelling glimpse of one of the ways in which ID speech facilitates word segmentation and possibly other aspects of language acquisition.

#### ACKNOWLEDGMENTS

This research was supported by the Waisman Center, the University of Wisconsin–Madison, National Institutes of Health Grant R01HD37466, National Science Foundation Grant BCS–9983630, and the University of Wisconsin–Madison Sophomore Summer Honors Apprenticeship. We thank Lindsey Dreyfus, Dana Emerson, Katrina Dvorak, Josh Kapfhamer, Stephanie Saeger, Becky Seibel, Katherine Stubler, and especially Kim Zinski for their assistance in conducting this research, as well as the parents of the infants for their time and generosity.

## REFERENCES

- Aslin, R. N. (2000, July). *Interpretation of infant listening times using the headturn preference technique*. Paper presented at the International Conference on Infant Studies, Brighton, England.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324.
- Cooper, R. P., & Aslin, R. N. (1990). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65*, 1663–1677.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Boston: MIT Press.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497–1510.
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Berkow (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). London: Oxford University Press.
- Fernald, A., & Cummings, A. E. (2003, April). *Ferguson's "clarification hypothesis" revisited: Does ID-speech facilitate word learning and word recognition by 18-month-olds*. Paper presented at the meeting of the Society for Research in Child Development, Tampa Bay, FL.
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279–293.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*, 209–221.
- Fernald, A., Pinto, J. P., Swingle, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*, 228–231.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development, 67*, 3192–3218.
- Frick, J. E., & Richards, J. E. (2001). Individual differences in infants' recognition of briefly presented visual stimuli. *Infancy, 2*, 331–352.
- Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11*, 43–79.
- Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language, 22*, 703–726.
- Hayes, J. R., & Clark, H. H. (1970). Experiments on the segmentation of an artificial speech analogue. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 221–234). New York: Wiley.
- Hertel, P. T., & Rude, S. S. (1991). Depressive deficits in memory: Focusing attention improves subsequent recall. *Journal of Experimental Psychology: General, 120*, 301–309.
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research, 5*, 69–95.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*, 548–567.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1–23.
- Kaplan, P. S., Bachorowski, J., & Smoski, M. J. (2001). Role of clinical diagnosis and medication use in effects of maternal depression on infant-directed speech. *Infancy, 2*, 537–548.
- Kaplan, P. S., Bachorowski, J., & Smoski, M. J. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychological Science, 13*, 268–271.

- Kaplan, P. S., Jung, P. C., Ryther, J. S., & Zarlengo-Strouse, P. (1996). Infant-directed vs. adult-directed speech as a signal for faces. *Developmental Psychology, 32*, 880–891.
- Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal  $f_0$  dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development, 67*, 205–217.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684–686.
- Liu, H., Kuhl, P. K., & Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*, F1–F10.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498–550.
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition, 7*, 127–145.
- Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2003). Infant visual recognition memory: Independent contributions of speed and attention. *Developmental Psychology, 39*, 563–571.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.
- Schieffelin, B. B. (1990). *The give and take of everyday life: Language socialization of Kaluli children*. Cambridge, England: Cambridge University Press.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3*, 365–394.
- Sluijter, A. M. C., van Heuven, V. J., & Pacily, J. J. A. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America, 100*, 2471–2485.
- Steedman, M. (1996). Phrasal intonation and the acquisition of syntax. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 331–342). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*, 188–195.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review, 9*, 335–340.
- Wagner, S. H., & Sakovits, L. J. (1986). A process analysis for infant visual and cross-modal recognition memory: Implications for an amodal code. *Advances in Infancy Research, 4*, 195–217, 240–245.
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development, 17*, 323–333.
- Venditti, J. J., Jun, S. A., & Beckman, M. E. (1996). Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 331–342). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

## APPENDIX

mo dibo lagoti nifopa kuda fa  
 mo nifopa lagoti kuda dibo fa  
 mo lagoti dibo kuda nifopa fa

mo kuda dibo nifopa lagoti fa  
mo dibo kuda lagoti nifopa fa  
mo lagoti dibo nifopa kuda fa  
mo kuda nifopa dibo lagoti fa  
mo nifopa lagoti dibo kuda fa  
mo dibo nifopa kuda lagoti fa  
mo lagoti kuda nifopa dibo fa  
mo kuda lagoti dibo nifopa fa  
mo nifopa dibo lagoti kuda fa