Prosodic differences in mothers' speech to toddlers in quiet and noisy environments

ROCHELLE S. NEWMAN
University of Maryland

ADDRESS FOR CORRESPONDENCE
Rochelle Newman, Department of Hearing & Speech Sciences, University of Maryland, College Park, MD 20742. E-mail: rnewman@hesp.umd.edu

ABSTRACT
We examined mothers' speech to 2-year-old children in both quiet and moderately noisy conditions. Mothers were recorded while teaching their children two words, one of which occurred in the context of other people speaking. Parents used characteristics of infant-directed speech (IDS) to these older children. Even more interesting was that many of the prosodic changes typical in IDS were accentuated in noise. In a context in which multiple people were speaking simultaneously, mothers spoke to their children with both increased pitch and increased word duration relative to when they spoke in a quiet condition. However, these changes were quite small, lending only mixed support to proposals that one of the advantages of IDS is that it is easier to separate from background noise.

One of the most fundamental facts about speech to children is that it is different from speech to adults. Not only do people tend to speak in shorter sentences with longer pauses when speaking to infants, but they also change the prosodic characteristics of their speech (Fernald & Simon, 1984; Masataka, 1992). Speech to young children tends to have higher pitch, greater pitch variability, and more prosodic repetition (Fernald et al., 1989; Shute & Wheldall, 1989). It also tends to be hyperarticulated, with more extreme vowel formant structure (Bernstein Ratner, 1984; Kuhl et al., 1997) and less overlap among the productions of intended consonants (Malsheen, 1980).

One result of this change in speech style is that it encourages infants to listen longer. Infants tend to prefer listening to speech with these characteristics, especially the pitch characteristics (Cooper, Abraham, Berman, & Staska, 1997; Fernald, 1985, 1992; Fernald & Kuhl, 1987; Masataka, 1996; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989; Werker, Pegg, & McLeod, 1994). In numerous studies, infants have been shown to listen longer to adults speaking in an infant-directed manner, especially by the time the infants reach 4 months

of age. They also are judged by adults to be more responsive overall when listening to this speech style (McLeod, 1993).

However, this speech register is not limited to use with infants, for whom these attentional biases have been demonstrated. Changes in speech style continue until children are into their school-aged years (Warren–Leubecker & Bohannon, 1984). Furthermore, similar speech styles are frequently used with the elderly and with foreign speakers (Caporael & Culbertson, 1986), people who may be more likely to dislike listening to infant-directed speech (IDS) than to prefer it (Ryan, Hamilton, & See, 1994; Whitbourne, Culgin, & Cassidy, 1995). These results suggest that speakers may not be using an IDS manner with a particular goal (or function) in mind.

What other consequences might the use of this speech style have? Some research suggests that children may be better able to detect clausal boundaries in IDS (Kemler Nelson, Hirsh–Pasek, Jusczyk, & Cassidy, 1989). Furthermore, Golinkoff and Alioto (1995) found that adult listeners learned new words better when spoken in an infant-directed manner. Fernald and Mazzie (1991) suggested that this speaking style may serve to highlight new words by giving them primary stress. These findings suggest that IDS can have linguistic, as well as attentional, benefits, particularly by highlighting relevant aspects of the signal.

IDS also may have affective advantages (McLeod, 1993). Werker and McLeod (1989) showed videotapes of infants to adult listeners. The adults rated those infants who were listening to IDS during the video recording as more appealing than those infants who were recorded while listening to adult-directed speech. This suggests that infants' greater attention to IDS may then alter adults' perception, helping to establish greater emotional rapport (Werker & McLeod, 1989).

Fernald (1984) suggested several other possible benefits of IDS, including the fact that the long glissandos may enhance perceptual coherence of the speech stream. However, these other suggestions have not received as much experimental support. Thus, there have been three main proposals for the functions of this style of speech (McLeod, 1993; Sachs, 1977): it highlights relevant aspects of the communication, making it easier for children to pick up on linguistic cues; it maintains children's attention; and it can increase positive emotional interactions between parents and children.

A proposal that has received somewhat less attention is that these changes in speech style may also make parental speech easier to separate from background noise (Colombo, Frick, Ryther, Coldren, & Mitchell, 1995; Fernald, 1984). Fernald (1984) suggested that the higher fundamental frequency would result in greater subjective loudness, increasing the signal to noise ratio of the signal. This would result in the IDS standing out against background noise. Colombo and colleagues (1995) also suggested that these types of stimuli would be more detectable in noise, and they tested infants' detection of sweep tones that matched either the typical intonational patterns of adult-directed speech or of IDS. They found that infants were better able to detect the sweep tones resembling IDS when presented in the context of white noise.

Moreover, many of the acoustic cues found in IDS are similar to cues shown to be important in adult streaming. For example, differences in pitch and pitch

variability are both important cues in musical stream segregation (Bregman, 1990; Brokx & Nooteboom, 1982), and differences in fundamental frequency are important for speech streaming (Darwin & Hukin, 2000). Furthermore, because infant-directed and adult-directed speech samples are likely to have less acoustic similarity with each other than would two samples of adult-directed speech, using IDS may ensure that the caretaker's voice is as dissimilar as possible from background speech (which is more likely to be adult directed).[1] This similarly raises the possibility that IDS would be easier to attend to than adult-directed speech in the context of noise. Newman and Weppelman (2002) tested this hypothesis with adult listeners. They presented listeners with the speech of two women speaking simultaneously and asked them to shadow a target voice. They found significantly better performance when the target voice spoke in an infant-directed manner than when she spoke in an adult-directed manner. This suggests that one consequence of IDS is that it serves to differentiate the speech of a caregiver from the speech of other people talking in the background (because that background speech is more likely to be adult directed).

Thus, IDS may be easier to attend to than adult-directed speech, particularly in noisy environments. If so, one possibility is that adults will therefore tend to make use of this speech register more often (or to a greater extent) when they are in these noisy environments. Because the acoustic cues that exemplify IDS appear to be helpful in the context of background noise, parents could potentially use these cues more often in just those environments in order to make the listening task easier for their children. This would imply that IDS would be more likely to be used in difficult listening situations or that those parents who regularly use IDS with their children might be expected to do so to a greater extent in noisier environments.

This clearly need not be the case. Even if IDS is particularly useful in noisy environments, that does not mean that parents would be sufficiently aware of the fact to make use of it (either purposefully or not). However, if parents are intending to communicate with their children, it makes sense that they would use whatever mechanisms were available to them to do so. Speakers do appear to make other sorts of acoustic alterations specifically for the listeners. For example, Gregory, Jurafsky, and Healy (2001) showed that speakers lengthen words specifically when they are new to the listener, regardless of the words' novelty to the speaker. Thus, there is some evidence to suggest that talkers regularly alter their productions in reference to the particular knowledge of the listener in that context.

An increase in IDS while in the presence of noise could also be made for attentional reasons, rather than to make the act of segregating speech easier. In noisy environments, children are more likely to be distracted by outside events. Caregivers may therefore need to try harder to capture their children's attention, and in so doing they may use more extreme versions of IDS. Children's feedback may shape the speech patterns that a parent adopts, and differential responsiveness to caregivers in noisy environments may encourage more extreme child-directed speech.

There have been findings suggesting that adults modify their speech to other adults when speaking in noise. When placed in extremely noisy environments

(>80-dB masking noise), adult speakers will increase their intensity (Lombard, 1911, as cited in Lane & Tranel, 1971), raise their fundamental frequency, and some will speak more slowly as well (Hanley & Steer, 1949; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). Likewise, both adults and children will make changes in their speech production when exposed to extreme auditory feedback, especially feedback consisting of multitalker babble. Rivers and Rastatter (1985) showed that both adults and children had more variable productions while speaking in multitalker noise presented at a level of 90 dB. They suggested that this was due to a lack of fundamental frequency (F0) control under these conditions, perhaps caused by competition between processing the incoming speech and formulating outgoing messages. Regardless of the cause, these findings clearly show that speakers do alter their speech in the context of noise. Yet many of these speech changes appear to be present only at fairly extreme noise levels, at least for adult listeners (Tartter, Gomes, & Litwin, 1993). Whether such changes might occur at lower noise levels with younger listeners has yet to be investigated.

Thus, for several reasons, we might expect parents to use a more extreme IDS (or child-directed speech) style in noisy environments. The present paper is an attempt to investigate this issue.

The procedure in this paper was modeled after Aslin and colleagues (Aslin, Woodward, LaMendola, & Bever, 1996). They asked parents to teach their 12-month-old infants two of three new words during the course of a single test session. The particular words were chosen such that they had onset boundaries that were potentially difficult to segment; this allowed the researchers to determine whether the parents made attempts to highlight these boundaries in any way. Mothers were tape-recorded using a lapel microphone; this recording was then used for the creation of transcripts, as well as for acoustic analysis.

We used a similar procedure, in which caregivers were asked to teach their children two new words. One of these was taught in a quiet environment and the other in a slightly noisy environment, in which other voices spoke in the background. The parents' speech to their children was recorded and compared across the two conditions.

We examined a number of properties that have been previously shown to be important in infant-directed speaking styles (Aslin et al., 1996; Brent & Siskind, 2001; Fernald & Kuhl, 1987; Fernald & Simon, 1984). One such property is the context in which the target word is placed. Brent and Siskind (2001) found that parents often present words in isolation when speaking to young children and that exposure to isolated tokens of a word predicts children's later acquisition of the word. Thus, the number of isolated occurrences is one feature of child-directed speech that parents might make greater use of in a more difficult listening environment. However, Aslin et al. (1996) found very different results from parents attempting to teach new words in a laboratory setting. They reported that, although some parents presented target words in isolation, others never did so; instead, mothers tended to place the target word in utterance-final position and tended to place emphatic stress on the target word (see also Fernald & Mazzie, 1991). Both final position and word stress appear to make the target

words more salient and extractable from the speech signal (Echols, 1996). In the present study, we examined all of these properties of parental speech: the number of times target words were presented in isolation, the proportion of time they occurred as the final word in an utterance, and the proportion of time they received emphatic stress.

In addition to these utterance-level properties, there are also prosodic cues that are characteristic of a child-directed speaking style. IDS typically consists of higher pitch and greater pitch variability (Fernald & Simon, 1984). Parents also tend to speak more slowly to young children (Drach, Kobashigawa, Pfuderer, & Slobin, 1969), which leads to elongation of the individual words. We examined each of these prosodic cues in mothers' speech in noisy and quiet conditions.

It is important to note that the noise level in the noisy condition was not so loud as to cause masking. The lesser distance between the mother and child in these sessions (compared to the distance between the child and the noise source) would have made the parental speech more intense than the background noise at the location at which the child sat. Thus, the question is not whether parents will choose to speak louder in order to raise the signal level above the noise. Rather, the question is whether parents will highlight their speech in ways that might make the separation of the two speech streams easier. Indeed, unless children are able to separate the different simultaneous streams of speech, increasing the relative amplitude levels will not help the child attend selectively to the target.

This study also allows us to examine characteristics of IDS with older children. Much of the research on child-directed speech has focused on speech spoken to very young infants (Cooper et al., 1997; Fernald & Simon, 1984; Fernald et al., 1989; Pegg et al., 1992; Werker & McLeod, 1989; Werker et al., 1994). There is evidence that parents continue to alter their speech until their children reach at least age 5 years (Warren–Leubecker & Bohannon, 1984). However, not all of the features of IDS have been examined with toddlers. Parents do appear to use the prosodic cues of IDS, such as fundamental frequency, with children in this age range (see Garnica, 1977), but they do not do so to the same extent as with younger children (Stern, Spieker, Barnett, & MacKain, 1983). Parents use a greater pitch range with 24-month-old children than with adult listeners, but they use an even greater pitch range with 4-month-olds. Similarly, although parents speak slower to children at 2 years of age than to adults, this difference is much less than that between 4-month-olds and adults (Stern et al., 1983). Thus, many of these prosodic modifications to children are beginning to decrease in extent by the time children reach 2 years of age. Nor is this merely a linear progression; parental speech to children is likely to undergo different types of clarification as a function of child language ability (Bernstein Ratner, 1984, 1996). Yet, a number of parental speech modifications (such as location of target words within a sentence) have not been examined across as many ages as have prosodic properties. Investigating whether the speech modifications used with 1-year-olds in earlier studies (Aslin et al., 1996) are also seen with toddlers is a secondary purpose of the present study.

METHODS

Participants

Twelve mother–child dyads participated in this study. All of the children (8 male, 4 female) were approximately 2 years of age (average age = 105.4 weeks; range = 103 weeks–109.9 weeks). Data from an additional two dyads were dropped from analysis as a result of experimenter error. Two of the children in the present study were twins; their mother was thus recorded twice. Thus, although there were 12 dyads, there were only 11 mothers. For analyses on children's accuracy, data from each of the 12 dyads were examined; for measures on parental speech, measurements taken with the twin who was tested second were excluded. However, analyzing these data by mothers or by dyads did not alter the results.

Procedure

Mothers were asked to teach two new words to their child, using a procedure based on that of Aslin and colleagues (Aslin et al., 1996). Their attempts were recorded onto audio tape using a Shure MX-185 lapel microphone connected to a Marantz portable tape deck.

We were concerned that parents might change the way they spoke to their children if they were aware their speech would be analyzed. We therefore told the mothers that we were interested in whether noise interfered with children's ability to learn new words. One reason we chose 24-month-old children (a relatively old age for studies on IDS) was because children of this age could be reasonably expected to learn new words, thus abetting our cover story. We stated that we were interested in recording any comments the child made while learning the words (e.g., if the child successfully repeated the word back), and that was the reason for using a microphone. We told the mothers that the reason we were asking them to wear the microphone, rather than their child, was that children tended to fiddle with the microphone if they wore it, causing static. Thus, having a microphone on the parent allowed us to record the child without the excess noise of having it actually on the child. At the end of the experiment, parents were informed of the true purpose of the study and were asked whether they suspected their voice was of interest. All mothers reported that they had had no such suspicions and indeed seemed quite surprised. The mothers were also asked at this time for permission to analyze their voice as part of the study, and all gave consent.

The mothers were asked to teach one of the new words in a quiet environment and the other in a noisy environment. In the Aslin et al. study (1996), mothers taught the two words in a single (combined) session lasting 5–10 min. Because we wanted the environment to differ across the two words, the two words needed to be taught separately from one another in the present study.

The two words we selected were *hedgehog* and *walrus*, and our props for these words were beanie-baby stuffed animals. These words were chosen because they potentially pose parsing problems for young children (see Aslin et

al., 1996). The toys were chosen to be of similar interest level to the children (they were of similar size and weight and had similar coloring) and because these words were unlikely to already be known by most children of this age. Furthermore, the two words are of similar length and stress pattern. Mothers were given 3 min to teach each word. Which animal occurred in which condition, which condition occurred first, and which animal occurred first were counterbalanced across dyads.

Mothers sat on the floor with their child in a laboratory test room. The noisy environment was created by having a computer (which was out of the children's line of sight) play a 10-min passage over a loudspeaker consisting of multiple people speaking among themselves as in a party environment. This recording had the natural variations in loudness level typical of these situations. The average noise level at the location of the room in which the parents sat was 60 dB SPL (±3). We allowed the computer to play the speech passage for a short period of time prior to the word-learning task, so that the children were comfortable in the situation; this prevented the children from being startled by the onset of the noise during the word-learning task. As the word-learning task only took 3 min, the passage never played completely through. The quiet environment was done in the same location, but without the computer playing this background speech.

We wanted the microphone to pick up the parent's speech but not to pick up much of the background noise. For this reason, we selected a cardiod lapel microphone. Parents attached this microphone to their clothing in a horizontal position, such that the microphone faced roughly in the direction of the parent's mouth but directly away from the computer playing the background speech. Thus, the background noise was in the shadow of the microphone. Because of this and the smaller distance between the mother's mouth and the microphone, the background speech tended to be at least 25 dB less intense on tape than the mother's voice. However, this method of recording did have one drawback, in that the amplitude of the mother's voice changed depending on whether her head was turned to one side or the other. Because the distance between the mother's mouth and the microphone was not constant, there was amplitude variability in recording caused by the direction in which her head was turned. This meant that we were unable to accurately analyze the mothers' average amplitude or amplitude variability across the 3-min passage, even though there is some indication in the literature that amplitude variability may increase when speaking to an infant (Fernald & Kuhl, 1987). Because mothers tended to turn and alter their position between utterances rather than within utterances, relative amplitude measures within utterances were still examined.

Mothers were given approximately 3 min to teach each of the new words. In case they ran out of things to say about these two animals, we gave them a "cheat sheet" of simple facts (such as what each animal liked to eat and where it lived). This was placed on the floor beside the mother so that she could refer to it if necessary. Although we tried to be as consistent as possible with regard to the time given for each word, we felt it best not to interrupt if the mother or child was in the middle of a statement. Thus, the amount of time per condition did not always last exactly 3 min. As our goal was simply to obtain a sufficient

sample of parental speech and parents did not necessarily spend equivalent proportions of time speaking, we were not overly concerned about this additional variation across mothers. Variation within mothers was controlled prior to analysis, however (see below).

Finally, we tested the children's word learning at the end of the session. The two target toy animals, along with several distractor toy animals, were placed in a row. (The distractor animals were bean-bag animals of similar size and material as the target animals.) The investigator asked the child to point to (or pick up) each of the animals in sequence. This was performed in a quiet listening environment only. The child's response to each request was recorded. Animals were placed back in the row following each question, and the two target items were never the first animal requested nor the last animal requested. If the child did not respond, the question was repeated; after the child did not respond on several prompts, the experimenter went on to the next animal. Animals were not repeated. Thus, children could select the named item correctly or incorrectly, but could only do so one time.

Coding

The audio recordings were digitized at a 44.1-kHz sampling rate with 16-bit quantization and recorded on computer disk. An investigator tabulated the number of times the mother said each of the target words and the location of each production (sentence-initial, sentence-final, sentence-medial, or in isolation); a second coder recoded these for reliability purposes (correlations on the number of tokens: total = 0.99, in isolation = 0.98, sentence-final = 0.92, sentence-medial = 0.97). Parents very rarely paused before or after target words unless they were at the beginning or ending of an utterance, so these location counts serve as a measure of how often the target words were clearly segmented in the speech signal. That is, words in the initial location and in isolation had clear onsets, whereas words in other locations did not; words in the final location and in isolation had clear offsets, whereas others did not. We did not count productions that occurred outside of the learning period (e.g., some mothers, while teaching the second animal, referred back to the first one as a referent; these productions were ignored). There was some indication that parents may have done this more often when the original teaching occurred in noise (and the new productions occurred in quiet; $n = 8$) than when the teaching originally occurred in quiet ($n = 3$), suggesting that parents may be aware that teaching in noise is less likely to have been successful. However, the number of instances is too small to make any definitive claim. We predicted that mothers would produce more tokens of the target word in the noise condition and that more of these tokens would be in salient locations (such as sentence-final) or in isolation. Each time that the target word was produced in a sentence (rather than in isolation), the investigator also recorded where the peak (or maximum) amplitude of the sentence occurred (i.e., whether it occurred in the target word itself or elsewhere in the sentence). This was examined by amplifying the digital signal until the sentence reached the maximum intensity the analysis program could accommodate. The location of this peak amplitude was then assessed by visual analysis

of the waveform. (All items were recoded by a second investigator; correlation = 0.87.) Although the location of the peak amplitude during a sentence could be influenced by head turns, most head turns and changes in position occurred between utterances rather than within them. Thus, these relative amplitude measures are less likely to be influenced by these variations. We predicted that mothers might be more likely to stress the target words in the noise condition. Although intensity differences are not the only cue to linguistic stress, they do serve as one such cue, which might be particularly relevant in the context of noise: the word in any phrase that is produced with the greatest intensity is going to be least likely to suffer masking from the background noise. Thus, we might expect that parents would be likely to give target words extra amplitude (relative to nontarget words) in the context of noise.

For these measurements, the exact amount of time spent teaching was an important factor, because a greater amount of time to teach a word would likely result in more productions. For this reason, we matched the two conditions on duration prior to tabulation. Thus, if the silence condition lasted 10 s longer in a given mother than did the noise condition, any speech occurring in the last 10 s of the longer condition was ignored.

We next examined the pitch of the mother's speech, predicting that mothers would use higher average pitch and greater pitch variability in the noise condition than in the silence condition. For pitch analyses, it was important to measure only the mother's speech. Thus, any time period in which the child spoke or when the only speech was that of the background talkers was removed from the analysis. Occasionally the child interrupted the mother, resulting in two voices speaking at the same time. These portions of the mother's speech were not included in the analysis. In addition, the mother sometimes turned such that the background speech was no longer in the shadow of the microphone and could thus be heard clearly. The background speech generally did not approach the loudness level of the mother herself, but any portions so affected were likewise removed from analysis. As we were interested only in the fundamental frequency, the speech was then downsampled to 10 kHz in order to make the analysis program's memory requirements less severe.

A spectral analysis program was used to measure the mothers' fundamental frequency at every 5 ms of the recordings. (The analysis routines used a fast Fourier transform to find the lowest major spectral peak; see Markel & Gray, 1976, for more details on the analysis routines.) From these successive measurements, two summary values were obtained: an overall mean pitch across the measures and the standard deviation of these measures (a measure of pitch variability). These pitch measurements were taken across the entire speech sample, rather than across just the target words, because we were predicting that parents might also place target words in different sentential locations in the context of noise. Because child-directed speech tends to involve exaggerated prosodic patterns that are repeated frequently (Fernald & Simon, 1984), we might expect that words at the ends of sentences would tend to have different pitch values on average than words in sentence-medial position. Rather than confound our pitch measures with measures of sentence position, we measured pitch across the entire condition instead of measuring it solely on the target words themselves.

Although there were potentially 3 min worth of each mother's speech per condition (and thus 36,000 pitch measurements for each mother and condition), no mother actually spoke for the entire 3 min. In addition to the substantial pauses and periods of time during which the child spoke, the mothers often produced voiceless consonants, which have no pitch. (Indeed, both the /s/ in walrus and the /h/ in hedgehog are voiceless.) Some mothers occasionally whispered as well (see Garnica, 1977, for similar findings). The analysis program indicated whenever it failed to find a pitch in a given time window. The pitch summary values were calculated only on those portions where the program successfully found a pitch. The number of measurements per mother per condition ranged from 5,488 for Mother 3 in silence to 18,429 for Mother 4 in silence.

One concern is that the background speech may have influenced the pitch measurements, despite the much greater signal strength of the mothers' speech. This could have altered the program's measurement of the mother's pitch. However, if the background speech was contributing to the measurements of pitch while the mother was talking, we would likewise expect that this background speech would contribute a pitch when the mother was producing voiceless consonants (such as /s/). That is, if the background speech was intense enough to cause a pitch perception by the program, this would have occurred both when the mother was producing voiced and voiceless items. This would therefore result in an overall greater number of pitch measurements in the noise condition than in the silence condition (because the program would successfully find a pitch during voiceless consonants in the noise condition but not during those voiceless consonants in the silence condition). This did not appear to be the case; the number of pitch measurements was quite similar across the mothers in the two conditions: for the noise condition, the average number of measurements was 9572 (range = 6756–16397); for the silence condition, the average number of measurements was 9412 (range = 5488–18429). This suggests that the presence of noise did not alter our frequency measurements. However, as a final check, we took the speech of one mother, which was originally recorded in quiet, and rerecorded it in both quiet and noisy environments.[2] These were then analyzed in the same manner as our original recordings. Because the two analyses were based on an identical recording, we would expect that the average pitch would be the same in the quiet and noise conditions if the presence of noise did not influence our measures. If the noise did influence our pitch measurements, we should find such an effect here. In fact, the average F0 for this recording was 219.98 Hz when recorded in the quiet condition and 219.87 Hz when recorded in the noise condition. Looking at the individual 5-ms measures for the two recordings, we found no significant difference between the measures taken in quiet and those in noise, $t (12,948) = −0.11$, $p > .05$. This is a clear indication that the presence of noise had no substantive effect on our F0 measurements.

For the final data from these pitch analyses, rather than match the two conditions on the amount of time the mother had to speak (as we did for the tabulations above), we matched them on the number of measurements per condition. Thus, if the analysis program found 6,500 measures for one condition but 7,000 measures for the second, our matched analysis examined only the first 6,500

measures for each. However, this actually had no effect on the results; the pattern of findings is identical whether we use the full recordings or only those portions matched for the number of measurements.

Our final analysis examined the duration of the target words in the two conditions. This is a major component to speaking rate changes, especially when an adult is attempting to speak clearly (Picheny, Durlach, & Braida, 1986). We predicted that mothers would produce the target words more slowly in the noise condition than in the silence condition. Here we selected only the first five tokens of each target in the two conditions. As would be expected from the prior literature (Bard & Anderson, 1994; Fowler & Housum, 1987), later repetitions of the target words tended to be much shorter than earlier repetitions in both the noise and silence conditions. This progressive shortening would not be a problem if parents repeated the words the same number of times in each condition. However, they did not, and we were concerned that this might influence the duration results if we used all tokens. For example, if a mother produced more tokens in the noise condition than in the silence condition (as predicted above), using the average duration across all tokens would result in the inclusion of more of these "shorter" versions of the target word in the noise condition than in the silence condition. This would make the average duration in that condition appear artificially short. Furthermore, we felt that the first few tokens were the ones in which the parent would be most concerned about producing the target word carefully; later on, the parents were often focusing on other new words (such as "tusks" for walrus), and the animal names received less emphasis. By focusing on the first five tokens of each target word, we should have an accurate representation of the duration of target words that the mother intends to be teaching. (The number 5 was selected because this allowed us to measure the same number of words across mothers; one mother only produced the target word five times in one of the two conditions.) We decided not to measure the overall number of words the parents used over the 3 min interval (another measure of speaking rate), because this would likely be confounded with the child's speech tendencies; in situations in which the child spent more time talking, the parent necessarily spent less. If children themselves are less likely to speak in noisy environments, this could artificially deflate measures of the parents' duration based on an overall measure of words per minute.

Because we had directional predictions for all analyses comparing quiet to noisy conditions, we used one-tailed paired t tests throughout our results. Although we did not explicitly control for the child's gender, this had no main effects and no interactions with any of the analyses and we therefore collapsed across gender in all of the analyses reported below. (However, this lack of a gender effect may be partly due to unequal samples, resulting in low statistical power.)

RESULTS

We first examined the children's success at learning the words in the two conditions. Accuracy here depends partly on definition: for example, when some

children were asked to point to the walrus, they did so correctly but then contin-
ued to point to that same object when asked to point to several other animals.
By an extremely strict accuracy criterion, where "correct" means that the chil-
dren both pointed to the correct animal when asked and never pointed to that
object when asked for something else, there were a total of six correct answers
to the target animals across all 12 children (out of 24 possibilities; no responses
and "I don't know" responses were counted as incorrect for this analysis). Three
of these correct answers were to the animal introduced in noise and three to the
animal introduced in silence.

However, this strict accuracy criterion may be an unreasonable expectation
for children of this age. By a less strict criterion, where we simply counted
whether the child selected the correct animal when the target items were named,
there were 9 correct answers to the target animals, again out of a possible 24
responses. This method of testing allows us to estimate an expected performance
based purely on chance. (It is not clear what "chance" performance is with the
stricter criterion.) Because there were five animals to choose among for each
question, we assume chance performance is 20%. Across the two target animals,
children averaged .75 correct responses (out of a maximum of 2); this is signifi-
cantly different from chance performance, $t(11) = 1.95$, $p < .05$.[3] This suggests
children did show some evidence of having learned the target items, although
this learning is not very robust.

Comparing across the two items, there was no difference in learning for the
animal introduced in noise as compared to the animal introduced in quiet, $t(11) =
1.15$, $p > .05$. Children did not appear to learn the words any better in the silence
condition than in the noise condition. Given the weak learning effects overall,
this lack of a difference is not terribly surprising. There was a trend towards an
order effect, in that most of the correct responses to the target words occurred
to the word taught first, $t(11) = 2.16$, $p = .054$, two tailed. Indeed, over half
the children correctly identified the animal they learned first, $t(11) = 2.58$, $p <
.05$. This may be an indication that children were more interested in the begin-
ning of the study than in the second part of the study. Looking only at the animal
learned first, there was a trend toward better performance for those children who
were taught that animal in noise than those taught that animal in silence, but
this was not significant, $t(10) = 1.86$, $p < .10$, two tailed. Overall, then, there
is no evidence of an influence of noise (or parent's differential teaching in noise)
on children's learning.

Our main focus in this paper was on the mother's speech, however. We first
asked the mothers if they had guessed the true nature of the study, and none of
them reported having done so. In fact, most evinced complete surprise that their
speech was of interest in the study. Although speech recorded in laboratory
settings may differ from speech outside the lab in many ways, we feel fairly
confident that the parents in the present study did not purposely try to alter their
speech in accordance with our experimental predictions.

We first examined the locations in which the target words were spoken. These
data are shown in Table 1. As was the case in Aslin et al.'s (Aslin, 1993;
Aslin et al., 1996) research, most mothers frequently placed the target words
in sentence-final position. Aslin and colleagues suggested that this location is

Table 1. *Frequency of occurrence of target words across conditions* (standard deviations)

	Noise	Silence
Mean number of		
Productions	26.0 (13.2)	25.4 (16.6)
Productions in isolation	5.0 (6.5)	5.2 (11.2)
Productions at end of sentence	11.5 (5.5)	10.1 (6.1)
Multiword utterances only		
Proportion at end	58.3 (20.0)	50.6 (24.0)
Proportion most intense	69.4 (14.1)	69.4 (24.5)

particularly salient for infants and that mothers might place words there in order to highlight them. This appears to also be the case with the older children tested here. Even this was not absolute, however: one mother in particular placed the target word in medial position over 75% of the time.[4] In fact, across all mothers, the proportion of tokens in sentence-final location, although high, was not nearly as high as in Aslin et al.'s study. They found that 89% of tokens (range = 76–100%) produced in multiword utterances in their study were in utterance-final position; in our study, the average was only 54% (combining across the noise and silence conditions) with a range from 24 to 100%. This suggests that the salient-location strategy may be beginning to be less important by the time children reach 2 years of age, although longitudinal studies would be needed to explore this directly. Another possibility is that different parents select different means of highlighting words for their young children; although placing target words in utterance-final position may be a strategy used by some parents, other parents choose other means to reach the same ends.

The extent to which parents spoke the target word in isolation varied between mothers; although some mothers did this quite frequently (over half of the time for one mother), others never used the word in isolation. This also did not predict children's accuracy in the final word test. For the word introduced in noise, the correlation between the number of occurrences in isolation and children's later accuracy was −0.10, $p > .05$; for the word introduced in silence, $r (12) = 0.50$, $p > .05$. Although Brent and Siskind (2001) have reported that exposure to words in isolation may facilitate vocabulary development, this did not appear to be a contributing factor in the present study.

Parents also frequently made the target word the most intense word in the sentence, according to acoustic measurements of peak amplitude. Across mothers, this occurred approximately two-thirds of the time, and all mothers did this at least one-third of the time (average = 69.4%, range = 40.3–93.9%). This is strikingly akin to the values for word emphasis reported by Aslin (1993) for 1-year-old infants, because he found that 68% of utterances had emphatic stress on the target word. This suggests that emphasis or increased amplitude of the target word is a fairly ubiquitous aspect of IDS, even with older children. Indeed, unlike the location of the target word in the sentence, this strategy of

highlighting the target word with emphatic stress does not appear to decrease as children age.

These results demonstrate that parents used speech patterns typical of child-directed speech, even though some of these changes were beginning to wane with the older children in the present study. However, parents did not appear to use these patterns differentially in the two conditions (noise and silence). There was no difference between conditions in the total number of times the target word was produced, t (10) = 0.20, p > .05; the proportion of time the word was produced in isolation, t (10) = 0.90, p > .05; or in the proportion of time the target was the most emphasized word in the sentence, t (10) = 0.01, p > .05. Parents were less likely to place the word in the middle of the sentence (where it would be hardest to segment) when the word was presented in noise, t (10) = 1.85, p < .05, suggesting that parents may be avoiding the most difficult word-learning situations while in a noisy environment. However, this was only significant with a directional test, suggesting that the effect is not very strong. Moreover, the mean length (in words) of the utterances containing the target word was no smaller in the noise condition than in the silence condition, t (10) = 0.21, p > .05; 4.25 words in noise and 4.29 words in silence. In general, there was only weak evidence to suggest that parents were altering their highlighting strategies when the children were in noisier environments.

We then proceeded to examine the prosodic characteristics of maternal speech. Here, we did find significant differences between conditions on several measures. Mothers consistently used higher pitch in the noise condition, t (10) = 3.68, p < .005. All but one of the mothers showed this pattern, the average pitch in the noise condition being approximately 218 Hz (11.6-Hz SD) and in silence being roughly 208 Hz (15.5-Hz SD; see Figure 1). Mothers also tended to produce the target words with a longer duration in the noise condition, t (10) = 1.87, p < .05, although this was significant only by the one-tailed analysis. On average, the target words were 727 ms in the noise condition (148-ms SD) but only 641 ms in the silence condition (109-ms SD). In contrast, although mothers did tend to use the high pitch variability indicative of IDS, this did not differ between conditions, t (10) = −0.93, p > .05.

DISCUSSION

Mothers in the present study clearly used several aspects of an IDS style to their 2-year-old children in both the quiet and noisy conditions. Not only did the mothers appear to use the extreme pitch variability typical of speech to infants, but they also tended to place target words in sentence-final position and to make these words more intense than other words in the sentence. Although the effect of word location within a sentence was not as strong as that found with 1-year-old children in prior studies, the percentage of time that the target words were emphasized was nearly identical to that with younger infants. This suggests that parents continue to adjust their speech to these older children in much the same way that they adjust their speech to young infants. Future work, using adult-directed speech as a comparison, will be needed to explore this issue more directly.
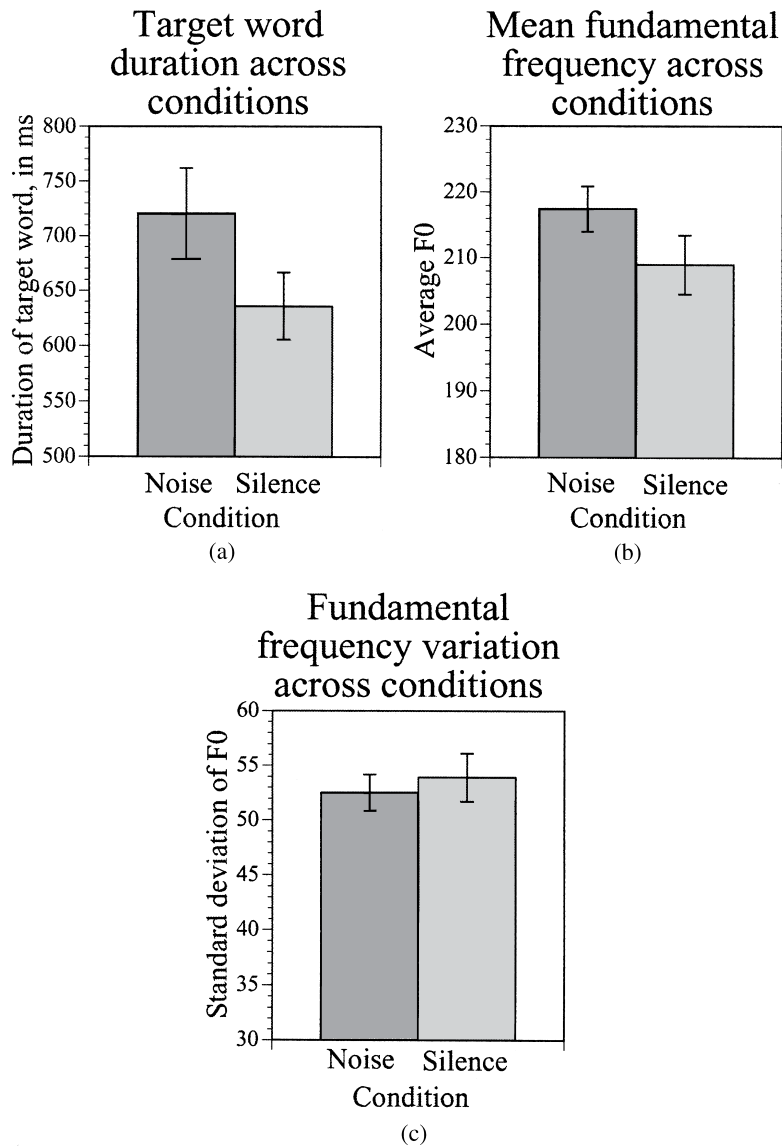
Figure 1. (a) The mothers' average pitch in noise versus silence conditions, (b) the average duration of target words in noise versus silence conditions, and (c) the standard deviations of mothers' pitch in noise versus silence conditions. F0, fundamental frequency.

The main focus in the present study, however, was not on whether parents used these cues in general, but whether they might use them differentially in different contexts. Here, we found mixed results. Although mothers did not change *what* they said in the noise condition (i.e., they did not repeat the target words more often or use different sentence structures in this condition), they did seem to change *how* they said it. Mothers spoke with both a higher pitch and at a slower speaking rate when there was other noise in the background. Speaking rate and voice pitch are two of the most often-reported changes in IDS relative to adult-directed speech (Fernald & Simon, 1984; Stern et al., 1983), and they appear to also be the changes most sensitive to context.

One of the other frequently noted changes in speech to infants is the variability in fundamental frequency, and this particular cue has been shown to be extremely important in young infant's preference for this speaking style (Fernald & Kuhl, 1987). Surprisingly, we did not see any changes in that measure between conditions in the present study. Why would mothers alter their average pitch without altering their pitch variability? We have no clear answer to this question, but there are a number of possibilities. One potential reason is related to ceiling effects. Mean pitch in the silence condition was not particularly high (less than 210 Hz on average), suggesting that there was a great deal of room for variation across contexts. Indeed, the average pitch values for the mothers in this study was similar to that found for typical adult-directed speech (Peterson & Barney, 1952), rather than IDS, suggesting that the mothers may not have been using the higher pitch values of IDS at all with these older children.[5] On the other hand, the frequency range in the silence condition was over 300 Hz; it may not have been possible for parents to increase this value still further for the noise condition.

Regardless of the reason for the lack of an effect in F0 variability, the finding that both duration and average pitch are increased in these noisy environments suggests that parents are sensitive to the listening environments their children are experiencing. These changes also provide support for the notion that one advantage of IDS is that it is easier to segment from adult-directed background speech. Indeed, increasing syllable duration is a primary change that speakers make when asked to speak more clearly (Picheny et al., 1986) and talkers who speak more slowly are judged as being more intelligible (Draegert, 1951), suggesting that the parents in the present study may be trying to compensate for the difficult listening conditions.[6] Future research will be needed to determine whether children benefit from these speech changes in the same way that adults have been shown to do.

However, the effect of pitch changes, although significant, is also quite small. On average, the parents differed by only 9 Hz across the two conditions. This certainly varied from parent to parent (two parents had differences of over 20 Hz, for example) and likely varied from moment to moment within the 3-min productions. Yet, on average, the difference in frequency was less than a semitone. It is not clear how large a pitch difference would be required to be noticeable. A 9-Hz difference would clearly be perceptible in a direct comparison: studies examining frequency difference limens (see Moore, 1989, p. 160) have found that frequency changes of less than 1 Hz are detectable by listeners for

frequencies in the region of the fundamental. Although these are results for pure tones rather than for speech, studies investigating speech have tended to find even smaller just noticeable differences, as low as 0.13% or 0.25 Hz (Rosen & Fourcin, 1986).

Despite this fine discrimination, listeners are unlikely to be sensitive to such small changes in fluent speech. Pierrehumbert (1979) suggests that a 12-Hz difference in pitch is needed for semantically meaningful distinctions (such as that of a difference in syllable stress). Some of our speakers did produce changes of this magnitude, but not all did. One possibility is that averaging across an entire 3 min is masking a distinction that is audible at some points in time; parents may increase their pitch at certain points in the sentence or on certain key words, but they may not do so consistently throughout the discourse. Another possibility is that smaller differences in pitch might still be useful for stream segregation, which is a very different task than detecting differences in stress. Still these results suggest that the average pitch difference in the current study may not be easily audible by listeners, at least for some talkers.

This fact, combined with the lack of significant findings for both nonprosodic features (such as word placement) and for pitch variability, suggests that these variations in parents' speech are unlikely to have a large influence on children's word learning. Although parents appear to be sensitive to the listening environment their children face, the alterations they made in their speech are unlikely to be large enough to be especially effective in facilitating comprehension in noisy environments. Why, then, are any such changes being made at all?

One possibility is that these speech changes are the result of a general attempt to speak more loudly or with greater vocal effort. When talkers increase their vocal effort, they do so by increasing their subglottal pressure and vocal-fold tension and by widening the opening of the vocal tract (Traunmüller & Eriksson, 2000). These changes result not only in greater amplitude but also in both longer vowel durations and increases in pitch, the two acoustic changes found in the present study (Bonnot & Chevrie–Muller, 1991; Rostolland, 1982; Traunmüller & Eriksson, 2000).[7] Thus, one possibility is that the changes seen here, rather than being the result of an increase in the use of an infant-directed speaking style, are by-products of an increase in vocal effort. This would explain why the increase in F0 was so small: if the increase in pitch was an unintended side effect of other changes intended to increase intelligibility in noise, there would be no reason to expect the change in pitch to be large enough to facilitate stream segregation. Note that this interpretation suggests that parents *were* modifying their speech in order to make it more intelligible in noise; however, the means with which they chose to do so was not related to an infant-directed speaking style per se.

Although an increase in vocal effort can explain the increases in pitch and duration found here, increases in vocal effort also tend to result in increased fundamental frequency variability (Rastatter & Rivers, 1983; Rivers & Rastatter, 1985), which we failed to find. Thus, the results from this study are not entirely in line with those predicted by the assumption of increased vocal effort. Future work will need to explore this vocal effort hypothesis more directly.

A related possibility is that the changes found in mothers' speech in the present study are not actually caused by parents' sensitivity to their listeners but are instead an effect of noise on talkers more generally. That is, increasing fundamental frequency and word duration in the context of noise might be something that speakers would do even when talking to themselves, as an automatic response to being in a noisy environment. As mentioned earlier, a number of studies have shown such effects of background noise on speech, at least when noise levels become extremely intense (Hanley & Steer, 1949; Lane & Tranel, 1971; Lombard, 1911; Van Summers et al., 1988). However, Tartter et al. (1993) found that, although duration changes did appear to be present even at lower noise levels, more like those examined here, fundamental frequency differences did not.[8] This gives us some reason to believe that the changes seen here are not simply another example of speech changes in noise.

Even if the speech changes found here are shown to occur in the presence of adult listeners at more moderate noise levels, this would not necessarily argue against the suggestion presented initially, that these speech changes could make separating a voice from the background easier for listeners. Indeed, if the acoustic changes in child-directed speech serve to make the target voice easier to follow, it stands to reason that talkers would use some of these same acoustic cues when speaking to anyone in a difficult listening environment. Even the changes in extreme levels of noise have been argued to occur primarily for the purpose of improving speech intelligibility. Listeners find that speech produced under these noisy conditions is easier to recognize, regardless of the noise level presented at test (Van Summers et al., 1988). Perhaps those changes necessary for adult listeners in extreme situations may be useful for young children even in the more moderate noise environments typically experienced in the home. The trend toward children's better learning in the noise condition supports this idea, but further research is needed to explore this in more depth.

However, other evidence suggests that the types of speech differences found here do not always lead to increases in intelligibility. Although word lengthening does seem to result in more intelligible speech (Picheny et al., 1986), overall sentence length does not (Bradlow, Torretta, & Pisoni, 1996). Moreover, Bradlow et al. (1996) found no correlation between average pitch and overall intelligibility. Picheny et al. (1986) found that, although there was a slight bias toward a higher F0 when an individual attempted to speak clearly, this difference was not substantial. Thus, whereas the longer word durations in the noisy condition may be a means of improving intelligibility overall rather than a means of improving segmentation per se, this is unlikely to be the case for fundamental frequency changes. Increases in F0 may serve to improve segregation of speech signals from noise, but they do not appear to improve overall intelligibility.

It is possible that the relatively minor changes seen here may be related to the relatively small amounts of noise we presented. In order to avoid having the noise itself influence our measures of parental fundamental frequency, we were restricted to using relatively quiet amounts of background noise. It is quite possible that parents would make greater modifications to their speech when placed in a truly noisy environment. Unfortunately, presenting more intense noise without it influencing the recording would require presenting the noise over head-

phones, and parents would then be unlikely to believe that their children's behavior, rather than their own speech, was the focus of investigation.

In conclusion, many of the changes indicative of speech directed to infants continue through at least 2 years of age. Even more interesting was that parents also seem to make changes in their speech production, depending on the listening environment. While in the presence of background noise, parents tend to both increase their pitch and decrease their speaking rate. They are also less likely to place target words in sentence medial position. These types of changes have been shown to help adult listeners follow speech in noisy environments, and thus may provide a similar benefit to children. However, the changes seen in the present study were relatively small and might therefore be insufficient to clarify the signal in any meaningful way. Future work is needed to examine whether these acoustic changes are sufficient to facilitate comprehension in difficult listening environments.

NOTES

1. This will vary depending on the particular situation. When a parent brings a young child out with them in public situations (such as to the grocery store, the park, etc.), the background speech in the environment is quite likely to be adult directed. In the home, background speech is likely to include sounds from the T.V. and radio as well as speech from other members of the family; this may include both speech directed to adults, and speech directed to older children (which may be more or less adult-directed in nature, depending on the age of the child in question). Day-care settings may be an exception to the notion that most background speech is adult directed. Unfortunately, there have been few studies examining the types of sounds infants typically hear in their environment. Van de Weijer (1998) is a notable exception.
2. We thank an anonymous reviewer for this suggestion.
3. We used a one-tailed test, because we were only interested in whether children selected the target item more often than chance, not whether they picked it less often than chance. It is worth noting, however, that this comparison assumes chance performance was 20%; given that the foil items may well have been familiar to the children, this may be an underestimate of chance performance.
4. It is worth noting that the investigator found this mother's speech to be most intellectually interesting, supporting the idea that this mother was speaking in a more adult-directed manner. This mother is also the only mother not to increase average pitch in the noisy condition.

5. Our subjective impression is that the mothers did intersperse their speech with portions that had the high pitch characteristics of IDS but also included portions with quite low pitch, such as when making the walrus talk. This resulted in extreme variability but not extreme averages.

6. Picheny et al. (1986) did not find significant increases in pitch, however, suggesting that our mothers were not simply speaking more clearly. Draegert (1951) did find that speakers judged to be more intelligible overall tended to have higher fundamental frequencies, but the difference was only 0.2 Hz and was not significant.

7. While all of these studies demonstrated increases in vowel duration with greater vocal effort, results for consonants varied among the different studies, sometimes showing no effect and sometimes showing a slight decrease in duration.

8. Of their two talkers, one showed no changes in F0 until the noise level reached 60 dB; the other showed a difference at a 35-dB signal to noise ratio, which then disappeared and even reversed at higher noise levels.

REFERENCES

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

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134). Mahwah, NJ: Erlbaum.

Bard, E. G., & Anderson, A. H. (1994). The unintelligibility of speech to children: Effects of referent availability. *Journal of Child Language, 21,* 623–648.

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language, 11,* 557–578.

Bernstein Ratner, N. (1996). From "signal to syntax": But what is the nature of the signal? In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 135–150). Mahwah, NJ: Erlbaum.

Bonnot, J.-F. P., & Chevrie–Muller, C. (1991). Some effects of shouted and whispered conditions on temporal organization. *Journal of Phonetics, 19,* 473–483.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic–phonetic talker characteristics. *Speech Communication, 20,* 255–272.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81,* B33–B44.

Brokx, J. P. L., & Nooteboom, S. G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics, 10,* 23–36.

Caporael, L. R., & Culbertson, G. H. (1986). Verbal response modes of baby talk and other speech at institutions for the aged. *Language & Communication, 6,* 99–112.

Colombo, J., Frick, J. E., Ryther, J. S., Coldren, J. T., & Mitchell, D. W. (1995). Infants' detection of analogs of "motherese" in noise. *Merrill–Palmer Quarterly, 41,* 104–113.

Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior & Development, 20,* 477–488.

Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *Journal of the Acoustical Society of America, 107,* 970–977.

Drach, K., Kobashigawa, B., Pfuderer, C., & Slobin, D. (1969). *The structure of linguistic input to children* (Working Paper No. 14). Berkeley, CA: University of California. Language–Behavior Research Laboratory.

Draegert, G. L. (1951). Relationships between voice variables and speech intelligibility in high level noise. *Speech Monographs, 18,* 272–278.

Echols, C. H. (1996). A role for stress in early speech segmentation. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 151–170). Mahwah, NJ: Erlbaum.

Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey, & R. Golinkoff (Eds.), *The origins and growth of communication*. Norwood, NJ: Ablex.

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

Fernald, A. (1992). Human maternal vocalization to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). New York: Oxford University Press.

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

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

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26,* 489–504.

Garnica, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*, (pp. 63–88). Cambridge: Cambridge University Press.

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.

Gregory, M. L., Jurafsky, D., & Healy, A. F. (2001, September). *The role of the hearer in durational shortening*. Paper presented at the Architectures and Mechanisms for Language Processing Meeting, Saarbrücken, Germany.

Hanley, T. D., & Steer, M. D. (1949). Effect of level of distracting noise upon speaking rate, duration and intensity. *Journal of Speech & Hearing Disorders, 14,* 363–368.

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

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277,* 684–686.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14,* 677–709.

Lombard, E. (1911). Le signe de l'élévation de la voix. *Annales Maladies Oreilles Larynx Nez Pharynx, 37,* 101–119.

Malsheen, B. J. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. In G. H. Yeni–Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception.* New York: Academic Press.

Markel, J. D., & Gray, Jr., A. H. (1976). *Linear prediction of speech.* Berlin: Springer Verlag.

Masataka, N. (1992). Motherese in a signed language. *Infant Behavior and Development, 15,* 453–460.

Masataka, N. (1996). Perception of motherese in a signed language by 6-month-old deaf infants. *Developmental Psychology, 32,* 874–879.

McLeod, P. J. (1993). What studies of communication with infants *ask* us about psychology: Babytalk and other speech registers. *Canadian Psychology, 34,* 282–292.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing* (3rd ed.). New York: Academic Press.

Newman, R. S., & Weppelman, T. L. (2002). *Infant-directed speech facilitates stream segregation.* Manuscript submitted for publication.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior & Development, 15,* 325–345.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America, 24,* 175–189.

Picheny, M. A. Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29,* 434–446.

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustical Society of America, 66,* 363–369.

Rastatter, M., & Rivers, C. (1983). The effects of short-term auditory masking on fundamental frequency variability. *The Journal of Auditory Research, 23,* 33–42.

Rivers, C., & Rastatter, M. P. (1985). The effects of multitalker and masker noise on fundamental frequency variability during spontaneous speech for children and adults. *The Journal of Auditory Research, 25,* 37–45.

Rosen, S., & Fourcin, A. (1986). Frequency selectivity and the perception of speech. In B. C. J. Moore (Eds), *Frequency selectivity in hearing.* London: Harcourt Brace Jovanovich.

Rostolland, D. (1982). Acoustic features of shouted voice. *Acustica, 50,* 118–125.

Ryan, E. B., Hamilton, J. M., & See, S. K. (1994). Patronizing the old: How do younger and older adults respond to baby talk in the nursing home? *International Journal of Aging & Human Development, 39,* 21–32.

Sachs, J. (1977). The adaptive significance of linguistic input to prelinguistic infants. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 51–61). Cambridge: Cambridge University Press.

Shute, B., & Wheldall, K. (1989). Pitch alterations in British motherese: Some preliminary acoustic data. *Journal of Child Language, 16,* 503–512.

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10,* 1–15.

Tartter, V. C., Gomes, H., & Litwin, E. (1993). Some acoustic effects of listening to noise on speech production. *Journal of the Acoustical Society of American, 94,* 2437–2440.

Traunmüller, H., & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *Journal of the Acoustical Society of America, 107,* 3438–3451.

van de Weijer, J. (1998). *Language input for word discovery.* Wageningen, The Netherlands: Ponsen & Loijien, bv.

Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustical and perceptual analyses. *Journal of the Acoustical Society of America, 84,* 917–928.

Warren–Leubecker, A., & Bohannon, J. N. (1984). Intonation patterns in child-directed speech: Mother/father differences. *Child Development, 55,* 1379–1385.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43,* 230–246.

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.

Whitbourne, S. K., Culgin, S., & Cassidy, E. (1995). Evaluation of infantilizing intonation and content of speech directed at the aged. *International Journal of Aging & Human Development, 41,* 109–116.