

Twenty-two-month-olds discriminate fluent from disfluent adult-directed speech

Melanie Soderstrom & James L. Morgan

Brown University

In press: *Developmental Science*

Corresponding Author:

Melanie Soderstrom, Ph.D.

Dept. of Cognitive & Linguistic Sciences

Brown University Box 1978

190 Thayer St.

Providence, RI 02912

Email: melsod@brown.edu

Abstract

Deviation of real speech from grammatical ideals due to disfluency and other speech errors presents potentially serious problems for the language learner. While infants may initially benefit from attending primarily or solely to infant-directed speech, which contains few grammatical errors, older infants may listen more to adult-directed speech. In a first experiment, Post-verbal infants preferred fluent speech to disfluent speech, while Pre-verbal infants showed no preference. In a second experiment, Post-verbal infants discriminated disfluent and fluent speech even when lexical information was removed, showing that they make use of prosodic properties of the speech stream to detect disfluency. Because disfluencies are highly correlated with grammatical errors, this sensitivity provides infants with a means of filtering ungrammaticality from their input.

Introduction

First language acquisition involves the interaction of input from the environment with the abilities and knowledge the infant brings to the task. An abiding question concerns how responsibility should be apportioned between these two factors. Arguments favoring innateness have often been predicated on the notion of the poverty of the stimulus (e.g. Baker, 1979; Chomsky 1962, 1965; Peters, 1972). This idea has two components: first, that the input underdetermines the target grammar, and, second, that the input available is degenerate – in Chomsky’s words consisting of “...fragments and deviant expressions of a variety of sorts” (1965, p . 201).

While data always logically underdetermine inductive generalizations, it is possible to ask the extent to which the input underspecifies the grammar of a language. Computational models have shown that certain so-called underdetermined facts about the grammar can be recovered from the input (e.g. Elman, 2005; Lewis & Elman, 2001). Furthermore, infants possess sophisticated statistical learning mechanisms that might allow them to find these properties in the input (e.g. Gomez, 2002; Saffran, Aslin, & Newport, 1996). The claim that the input is degenerate has also been disputed by analyses of maternal speech that show it to be, if not perfect, at least generally well-formed (e.g. Newport, Gleitman, & Gleitman, 1977; Soderstrom, Blossom, Morgan, & Foygel, under revision). The high rate of disfluencies (hesitations, false starts, repetitions, and repairs) – and, hence, ill-formed utterances - in adult-directed (AD) speech (e.g. Nakatani & Hirschberg, 1994; Clark & Fox Tree, 2002) is not found in maternal, infant-directed (ID) speech.

To obtain well-formed input, therefore, young infants might attend primarily or exclusively to ID speech. Indeed, across a number of studies, young infants learning a number of languages show a robust preference for infant-directed speech produced by both men and women (Cooper, Abraham, Berman, & Staska, 1997; Cooper & Aslin, 1990; Fernald, 1985; Fernald & Kuhl, 1987; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989). This preference is found even for languages unfamiliar to the infant (Werker, Pegg, & McLeod, 1994). A small number of studies have also shown that infants are more successful at speech processing tasks when ID speech is used (Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Thiessen, Hill, & Saffran, 2005)

Nevertheless, accounts of acquisition cannot assume that infants learn solely from infant-direct speech. Not all infants are necessarily exposed to such speech. For example, Kaluli parents reportedly do not use a special register when speaking with infants and children the way Western parents do (Schieffelin, 1985). Moreover, some of the properties of ID speech make it less than ideal teaching material. ID speech is not only well-formed, but also syntactically simple. Maternal mean length of utterance (MLU) at a variety of infant ages has been found to be around 4 words – no more than half that of normal adult-directed speech. The vast majority of ID utterances are single clause utterances, and contain pronoun subjects (e.g. Broen, 1972; Morgan, 1986; Soderstrom et al., under revision). In fact, maternal speech may be highly well-formed *because* it comes packaged in short utterances that leave little opportunity for disfluency to manifest, and/or because shorter utterances are easier to produce. Ultimately, it would be beneficial for the language learner to access the full range of linguistic input, such as conjunction and embedding of clauses, and even full noun phrases in subject position, which in ID

speech are highly underrepresented. The potential benefit in attending to AD input would increase as the infant becomes competent in the simpler structures of the language.

Generally, it has been assumed that older infants and even young children attend more to ID than AD speech, but most studies testing these preferences have focused on determining how young the preference can be found (e.g. Pegg et al., 1992). The oldest age group tested in any study to date has been 10-14 months, and conflicting results have been found, with one study finding a preference for ID speech (Hayashi, Tamekawa, & Kiritani, 2001), and another finding no preference at this age (Newman, 2006). Furthermore, even if infants (of any age) attend preferentially to ID speech, this does not mean that AD speech is ignored by them entirely. One recent comprehensive analysis of the total input to a Dutch infant (van de Weijer, 2002) found that only 15% of the speech heard by the infant was directed to that infant - an additional 30% was directed at an older sibling. Therefore, a very substantial proportion of the speech heard by any infant may be adult-directed.

The ability (and need) to handle input more complex than ID speech may come relatively early. Infants as young as 16-18 months are beginning to demonstrate grammatical knowledge such as the basic word-order structure of the language (Hirsh-Pasek & Golinkoff, 1996) and dependency relationships between highly frequent function words and inflections (Santelmann & Jusczyk, 1998; Soderstrom, White, & Conwell, 2005). These findings suggest a possible role for the less fluent, but potentially more informative, input provided by adult-directed speech even before 2 years of age. This idea becomes more plausible if it can be shown that infants differentiate well-formed and ill-formed utterances in the absence of grammatical knowledge. In this

article, we will examine whether infants are perceptually able to distinguish well-formed from ill-formed fluent speech utterances.

Disfluencies often result from grammatical errors, as speakers backtrack to correct inappropriate morphology, missing words, and so forth. However, disfluencies may also *create* grammatical errors – in particular, by creating word strings or groupings of words that do not correspond to syntactic constituents. One subtle source of misinformation about grammatical structure arises from prosodic characteristics like pauses and intonational or rhythmic changes associated with disfluencies (Aylett, 2005; Bartkova, 2005; Loevgren & van Doorn, 2005; Nakatani & Hirschberg, 1994). These prosodic factors may create groupings of words that infants might confuse with the normal groupings based on the prosodic structure of fluent speech. For example, the disfluent pause (indicated by the # symbol) after “*we*” in the utterance “Yeah, so we # ended up...” creates the prosodic grouping “so we”, which is not a well-formed syntactic constituent. Such prosodic disfluencies are potentially problematic for infants, because well-formed prosodic word groupings may be used by infants to find syntactic word groupings – and syntactic dependencies among words within such groupings - in speech (e.g. Gleitman & Wanner, 1982; Morgan & Newport, 1981; Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000; Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003).

Adults can in principle identify and compensate for disfluencies during speech processing using lexical and grammatical properties of the word strings, such as incomplete words, ungrammatical word strings, or overt markers of disfluencies such as *uh* or *um* (Clark & Fox-Tree, 2002). However, the language learner by definition has an immature grammar, and must rely on information external to that system. One possible

source of information to infants about the well-formedness of utterances is their prosodic properties. If infants are sensitive to the *differences* between fluent and disfluent prosodic structure, they could use this to identify well-formed input, or exclude ill-formed input.

Lickley & Bard (1998) showed that adults are able to judge the locations of disfluencies in speech rapidly and accurately, without access to lexical or grammatical information. Participants were played speech and asked to judge whether disfluencies were present. The stimuli were “gated” so each speech sample was played repeatedly with successive word-level incrementing up to and past the disfluency point. Although the participants were not able to detect disfluencies when the sample ended before the *interruption point* (the point at which the disfluency occurred), they were able to correctly detect the disfluency immediately *at* the interruption point, before any additional words were uttered that might cue the listener that a disfluency had occurred. This result suggests that adults can use *prosodic* information at the interruption point to detect disfluencies. Given that infants are highly sensitive to the prosodic characteristics of speech (see Jusczyk, 1997, for a review), they may also be able to use such information to differentiate fluent from disfluent utterances.

Disfluent utterances have several characteristic prosodic properties. For example, words immediately following a disfluent interruption point tend to be stressed or hyperarticulated, particularly if they constitute a repair (Levelt & Cutler, 1983). Repeated words in disfluency have flat intonation, and tend to be of shorter duration, even if they are not fragments (Bartkova, 2005). There is also some evidence that the degree of pitch reset (increase in f_0) and increase in amplitude is greater after disfluent interruptions than after fluent prosodic boundaries within an utterance (Aylett, 2005; Nakatani &

Hirschberg, 1994). Cole et al. (2005) have also described a phenomenon of “prosodic parallelism”, whereby the reparandum (the portion containing an error that needs to be fixed) and the repair (the repeated portion in which the error is fixed) share prosodic characteristics. Prosodic characteristics are present, therefore, which might serve to differentiate well-formed utterances and well-formed prosodic boundaries, from disfluent ones. Here we examine whether infants are generally sensitive to the differences between fluent and disfluent speech, and specifically whether infants are sensitive to the *prosodic* differences between fluent and disfluent speech.

Experiment 1: Discriminating fluent and disfluent speech

In Experiment 1, we asked whether infants differentiate between disfluent and well-formed utterances in AD speech. Two ages of infants were tested – 20 to 23-month-olds, an age at which infants are already demonstrating some initial grammatical knowledge (henceforth the Post-verbal group), and 10-month-olds, an age at which infants are not yet demonstrating any grammatical knowledge (henceforth the Pre-verbal group). We chose 10 months as our Pre-verbal group because this is just prior to the youngest age at which there is evidence for sensitivity to the grammatical properties of language (e.g. Shady, 1996; Shafer, Shucard, Shucard, & Gerken, 1998).

If infants respond to highly salient, universal prosodic cues to disfluency, one might expect that both age groups would succeed at the task. Even very young infants use prosodic cues in analyzing speech input (e.g. Mandel, Jusczyk, & Kemler Nelson, 1994; Mandel, Kemler Nelson, & Jusczyk, 1996; Nazzi et al., 2000; Soderstrom et al., 2003). On the other hand, the prosodic characteristics of AD speech are less exaggerated than ID

speech (e.g. Fernald et al., 1989), and may be less salient to younger ages. Older, Post-verbal infants may be better equipped to detect such subtle cues at an age where they are most critical to the infant.

In this first experiment, infants were played paired recordings of normal English fluent or disfluent speech. We were interested in two related aspects of the infants' listening behavior. Our primary concern was whether they discriminate the fluent and disfluent recordings. If infants discriminate fluent and disfluent speech, it was predicted that they would show differences in their attention to the different recording types. We were secondarily interested in whether infants might show a preference, measured by *greater* attention, for the more well-formed recordings. To the extent that our experimental contexts mimic those of the real world language-learning environment, such an attentional bias for the well-formed speech would be in the best interest of a language learner seeking well-formed input.

Transcript analysis

The stimuli used in the experiment were based on a recording that was part of a larger study of infant-directed speech (Soderstrom et al., under revision; Soderstrom & Morgan, 2005). In this study, mothers were given recording equipment over a period of several months, and asked to record weekly their normal verbal interactions with their infants. The recording of interest was taken while mother, father and 2.5 month old infant waited in a crowded, noisy airport terminal lobby. There was a mix of adult- and infant-directed utterances in the 33 minute sample. The recording was transcribed by a trained transcriber using the CHAT system (MacWhinney, 2000), and checked by a second transcriber. Each utterance was also coded for significant fluent and disfluent prosodic

breaks (salient pausing, lengthening or pitch reset that would normally be marked by a comma, #, or disfluency marker in the CHAT system) within the utterance. Differences of opinion were resolved by discussion between the transcriber and the checker.

Table 1 shows the number of utterance-internal disfluent and fluent prosodic breaks in the sample based on transcriber coding. Two things are notable. First, there were a large number of disfluent prosodic breaks in the sample – more than a quarter of the utterance-internal breaks transcribed were disfluencies. Of course, the ambient noise level was high, and the mother chatted animatedly with a series of strangers as they waited for their flight – these factors may have served to increase the total number of disfluencies compared to a home environment. Second, although the numbers of utterance-internal fluent prosodic boundaries were comparable across speech addressee types, the rate of disfluencies varied enormously across addressees. In speech to adults, almost 40% of all utterance-internal prosodic breaks were disfluent, whereas in speech to the infant, only 2% were; only 3 of 96 disfluencies were in infant-directed utterances. This confirms that while AD speech may contain many disfluencies, ID speech is largely immune to the difficulties associated with disfluency in speech input.

Table 2 shows the numbers of all grammatically well-formed and ill-formed multi-word phrases and clauses preceding fluent and disfluent prosodic breaks, including disfluent utterance endings. This tally excludes prosodically isolated single words, for which grammaticality cannot be defined. About 80% of the word sequences preceding disfluent prosodic boundaries were syntactically ill-formed. Most of these word sequences were ill-formed because they were missing important constituents (e.g. “we just ...[verb phrase]”, “that’s the ...[noun]” “so we ...[verb phrase]”) By contrast, 99%

of word sequences preceding fluent prosodic breaks were grammatically well-formed. Of the six ungrammaticalities preceding a fluent-sounding prosodic break, four directly followed a disfluency, e.g., “this is a really strange uh # gate down here” (where “gate down here” was tallied as an ungrammatical prosodic word sequence preceding a fluent, well-formed utterance boundary). Of the remaining two, one was formed by the insertion of “you know” in the middle of a phrase: “...which is just, you know, right next to it..”. The other was preceded by an “um”, a non-prosodic cue to disfluency, but simply did not sound disfluent to the ear. Based on our transcribers’ judgments, the mother’s adult-directed speech in this sample, though often disfluent, also contains highly reliable prosodic cues to the ungrammaticality these disfluencies caused. With two possible exceptions, therefore, *all* of the ungrammatical word sequences were marked prosodically by disfluency. In the study reported here, we asked whether infants could discriminate fluent utterances from disfluent utterances modeled on those in the transcript.

Method

Participants.

The data included were from 24 infants between 20 and 23 months (range: 607-727, mean: 673 days, 12 male and 12 female), and 24 infants 10 months of age (range: 309-339 days, mean: 328 days, 11 male, 13 female). Five additional infants from the older age group participated in the study but their data were discarded due to fussiness or squirminess. No 10-month-old data were discarded in this experiment. All participants were normally developing infants with normal hearing from Providence, RI, USA, and had parents and caregivers who were native speakers of American English.

Stimuli.

Disfluent utterances were culled from the transcript to obtain real-world examples of disfluencies. These stimuli were used as models of disfluencies rather than attempting to create disfluencies in a laboratory environment because it was necessary to have a high density of disfluencies, which would be very difficult to recreate in a laboratory environment, particularly given that spontaneity, noise, and conversational interaction are important factors in creating an environment in which disfluencies are high. Because of the ambient noise in the recording, and to produce fluent matched controls, the original recordings were not used during the testing sessions. Instead, a trained speaker listened carefully to each utterance, and mimicked as closely as possible the prosodic characteristics of the original recordings. Fluent versions were created by modifying the disfluent utterances as little as possible (both syntactically and prosodically) to make them fluent. These control utterances were produced at the same time, and matched as closely as possible the prosodic characteristics other than those of the disfluency itself.

In a few cases, the utterance boundaries of the stimuli did not match up exactly with the original transcription. Some utterances were truncated from the original transcription at salient silent pause boundaries. This was done because the original speaker produced very long, run-on sentences that were originally transcribed as single utterances. Other utterances were combined from two smaller sentences or fragments, to preserve the salience and flow of the disfluency. The resulting utterances sounded natural as utterances, and did not sound more disfluent than the originally transcribed utterance divisions. Below is an example of a disfluent utterance and the fluent control version (# indicates pause, [///] indicates restart, +//. indicates self-interruption).

- (1) yeah , so we # ended up [///] we drove through Boston for like an hour, literally, with all the traffic +//.
- (2) yeah, so we ended up driving through Boston for like an hour, literally, with all the traffic.

Low-pass filtered versions of 40 fluent and disfluent utterances (filtered at 400 Hz, with 100 Hz of smoothing, using Praat software (Boersma, 2001)) were then rated by 8 adult listeners, using a 7 point rating scale, who judged how fluent the stimuli were, with 1 most fluent, and 7 most disfluent. Low-pass filtering speech removes the high frequency information necessary for recognizing individual words, but leaves the general prosodic structure intact. The raters were able to discriminate the fluent and disfluent samples with a high degree of accuracy, $t(7)=7.27$, $p < .001$. The average rating for disfluent utterances was 4.5 and for fluent utterances 3.3.

A subset of 20 utterance pairs with an average rating difference of at least 1 point and with similar durations was chosen. Nineteen of these selected utterances were adult-directed, and one was infant-directed. Using the PSOLA method in Praat, the durations of the fluent and disfluent versions of each utterance pair were equalized by slowing down the shorter version and speeding up the longer by equivalent amounts, while preserving the overall pitch characteristics. For 11 of these utterance pairs, the fluent version was shorter than the disfluent version, while for 9 utterance pairs, the reverse was true. Therefore, acoustic differences in the stimuli caused by the PSOLA modification could not be responsible for differences in infant listening preferences. A second set of 8 adult raters then verified that post-adjustment the disfluent versions were still highly discriminable from the fluent versions (again using low-pass filtering). The

raters were again able to discriminate the fluent and disfluent samples with a high degree of accuracy, $t(7)=5.69$, $p = .001$. The average rating for disfluent utterances was 4.7 and for fluent utterances 2.9.

A third set of 8 raters judged how closely low-pass filtered versions of these final stimuli matched the original speech samples in their prosodic characteristics, with 1 denoting very closely matched and 7 not at all matched. The disfluent versions were given an average matching score of 2.3, while the fluent versions were given an average matching score of 4.2 (to the original disfluent recording). Additionally, the pitch tracks of original and laboratory recordings of disfluent utterances were converted to semitone scores (in 50 ms bins) based on the average pitch of each utterance and compared. The average Pearson correlation between original and laboratory version was .62, which is high given that small timing differences would have large effects on this score.

The selected utterances were grouped into 4 disfluent and 4 fluent "passages" of 5 utterances each. One second pauses separated the utterances within each passage.

The full list of selected stimuli can be found in the Appendix.¹ Table 3 provides information about the properties of these disfluent utterances and their fluent matched controls. There were no significant differences between utterance types in duration, number of syllables, or number of repeated words – control sentences were constructed to minimize such differences. The disfluent passages did contain significantly more utterance-internal prosodic boundaries than the fluent passages, $t(19)=4.50$, $p < .001$. Along with prosodic disruptions, the four disfluent passages contained a total of 5 um/uhs and 9 part-words, which may also serve as signals to disfluency.

¹ Audio recordings of the stimuli used are also available for examination by contacting the authors. A subset are available online at <http://www.cog.brown.edu/~morgan/disfluency>

Procedure.

Infants were tested using the Headturn Preference Procedure (HPP, Kemler Nelson et al., 1995). Infants were seated in a testing booth on a caregiver's lap. The caregiver wore aviator's headphones that played music to mask the experimental speech sounds. The experimenter was situated in an adjacent room and monitored the infant's behavior on a video monitor. Each trial was initiated by a flashing light in the center of the front panel of the booth, to direct the infant's attention forward. When the infant was looking forward, the experimenter would signal the computer to extinguish the light and begin the trial.

For the Pre-verbal infants, a light would then begin to flash on either the right or left side of the booth. The side was randomized for each trial. When the infant oriented to this light, the experimenter would signal the computer by clicking and holding the mouse, and speech would begin to play from a speaker located behind the light. Speech would continue to play until either the passage was complete (approximately 21 s), or the infant looked away from the light for 2 consecutive seconds. The direction of the infant's eyegaze was recorded by the computer based on the experimenter's mouse presses. Speech continued to play during the trial while the infant was oriented away from the light, until the 2 s criterion was reached and the trial ended. If the infant did not orient toward the light for a minimum of 2 s during a given trial, the data were discarded and the trial was repeated. The dependent measure was the total amount of time each infant was oriented toward the light while the speech was playing, averaged across the four trials for each condition, fluent and disfluent.

For the Post-verbal infants, the procedure was essentially the same. However, due to the difficulty in keeping a 20-month-old seated and engaged for the length of the procedure, the apparatus was modified to use side-by-side TV displays. After the infant oriented toward the center flashing light, one of the screens would begin to flash a round orange circle. When the infant oriented toward this circle, the screen would display a black-and-white checkerboard form, and the speech would begin to play. The “video HPP” procedure was otherwise identical – the dependent measure was the infant’s orientation time toward the checkerboard form while the speech was playing.

Each testing session began with two pretest trials to orient the infant to the contingency between the lights (and video, for Post-verbal infants) and the speech stimuli. The pretest stimuli were taken from a previous study (Soderstrom et al., 2003) and chosen so as not to bias the infants toward listening to fluent-sounding or disfluent speech. The pretest trials lasted a maximum of 10 seconds, and consisted of repetitions of the fluent phrase “people by the hole” (extracted from the sentence “Today, *people by the hole* seem scary”) or the ill-formed “people buy the whole” (extracted from the sentence “In fact, some *people buy the whole* supply of them”). After these two pretest trials, infants were immediately tested with the 8 test trials presented in random order.

Results and Discussion

Means and standard errors of looking times are shown in Figure 1. The Pre-verbal infants showed no preference for either passage type. Mean looking time to the disfluent passages was 7.3 s ($SD = 2.82$) and to the fluent passages 7.2 s ($SD = 2.63$), $t(23) < 1$, $p > .5$, $d = .045$. Of the 24 infants in this condition, 7 infants preferred the fluent passages on average by more than 1 s, and 7 preferred the disfluent passages by

more than 1 s. By contrast, the Post-verbal infants preferred the fluent passages, looking to the fluent passages for an average of 7.6 s ($SD = 3.19$) compared with 6.2 s ($SD = 2.57$) for the disfluent passages. This was significant by 2-tailed paired t-test, $t(23) = 2.3$, $p < .05$, $d = .42$. Of the 24 infants in this condition, 13 infants preferred the fluent passages on average by more than 1 s, while only 3 preferred the disfluent passages by more than 1 s. A mixed ANOVA with age and passage type as factors found a marginal interaction between age and passage type, $F(1, 46) = 3.69$, $p = .06$.

Insert Figure 1 about here

For the Post-verbal infants, both hypotheses were confirmed. Infants showed differences in their looking times to the two passage types, and this difference was in the direction predicted by a preference for well-formed speech. Infants preferred to listen to the fluent speech samples. This finding suggests that by the time they are beginning to form a grammar, infants are able to distinguish fluent, well-formed utterances from disfluent utterances in adult-directed speech. Since all but one of the speech samples were based on adult-directed spontaneous speech, it also provides some evidence that Post-verbal infants do process adult-directed speech, at least under some circumstances. Before they reach the age of two, infants may be able to make use of adult-directed speech as input, despite its relative complexity and disfluency.

However, for the Pre-verbal infants, we found no evidence that they were differentiating fluent and disfluent speech. To verify that procedural differences had not caused the difference between the age groups, we tested an additional set of sixteen 10-month-olds in the video paradigm. Again, no preference was found. Mean looking time to the disfluent passages was 6.8 s and to the fluent passages 6.5 s, $t(15) < 1$, $p > .5$.

Three infants preferred the fluent passages on average by more than 1 s, and three preferred the disfluent passages. A mixed ANOVA with age and passage type as factors again found a marginal interaction between age and passage type, $F(1, 39) = 3.18, p = .08$. Despite previous work showing that infants are highly sensitive to prosodic characteristics of speech at a young age, Pre-verbal infants did not make use of prosodic information to differentiate the fluent and disfluent speech samples. It may be that infants at this age are not sensitive to the particular prosodic properties that indicate disfluency. Or Pre-verbal infants might attend less to adult-directed speech than older infants, although the similarity in looking times across the two groups argues against this possibility. A more likely possibility is that the cues that indicate disfluency are more subtle in adult-directed speech than infant-directed speech, and Pre-verbal infants have greater difficulty detecting them. Overall, AD speech is less exaggerated prosodically than ID speech – it may be that the specific prosodic characteristics that differentiate fluent from disfluent prosodic breaks are also less exaggerated.²

Experiment 1 demonstrates that Post-verbal infants can differentiate disfluent and fluent speech. This finding suggests that infants at this age may be able to access well-formed adult-directed input. However, this result does not establish whether infants are using prosodic information to succeed at the task. Post-verbal infants are by definition beginning to form grammatical knowledge of their language. It is possible that infants are relying on this knowledge, and not prosodic cues, to differentiate the passage types.

² Some pilot work in our lab may provide support for this latter possibility. We tested Pre-verbal infants on disfluent samples containing infant- and child-directed speech. We found a preference for the *disfluent* passages, suggesting that infants might differentiate fluent and disfluent passages in ID speech. However, the effect was not reliable across experimental conditions.

In Experiment 2, we examine whether prosodic information plays a role in the infants' ability to discriminate the speech samples, by excluding lexical and grammatical cues from the stimuli.

Experiment 2: Discriminating modified speech samples

Experiment 1 established that Post-verbal infants discriminate disfluent from fluent speech but did not identify the cues that infants are using to make this discrimination. In Experiment 2, the speech samples from the first experiment were phonemically modified to remove lexical and grammatical information. If infants are depending on lexical and grammatical information to detect disfluency, they should not be able to discriminate the modified samples. However, if they are using prosodic cues, which were preserved in the modification, they should continue to discriminate the samples.

A standard way to remove lexical and grammatical information, but not prosody, is to low-pass filter speech, as we did in Experiment 1 to obtain adult ratings of our stimuli. In pilot work, we ran an experiment with 25 22-month-olds in which the speech was low-pass filtered rather than phonemically modified. In this condition, infants did not show differences in their listening times to the fluent and disfluent passages. We suspect that the removal of phonological context made it more difficult to determine utterance boundaries, which may be important for detecting disfluency. Adults also anecdotally found it much more difficult to discriminate the samples when the concatenated passages were played in the testing room, although they easily discriminated individual low-pass filtered versions in our rating task. While a number of studies have been successfully run

with infants using low-pass filtered speech, these have generally tested much younger infants and have examined larger scale prosodic features, such as those differentiating rhythmic classes (Mehler et al., 1988 – newborn and 2-month-olds), or driving preference for maternal speech (Spence & Freeman, 1996 – newborns). One possible exception is a study demonstrating that 9-month-olds are sensitive to the placement of pauses at phrase boundaries in low-pass filtered speech (Jusczyk et al., 1992). However, the current study examines whether infants differentiate between different kinds or number of prosodic boundaries, not the location of these boundaries with respect to artificial pauses, and may therefore present a more subtle task.

We therefore chose a method that would isolate the prosodic characteristics of the samples from lexical and grammatical information, while preserving general phonological properties. Ramus and Mehler (1999) developed a method to artificially reduce the phonemes in a speech sample to a set of manner changes, which are then overlaid on the original pitch track. This method has been successfully used to test phonological and prosodic sensitivities in infants, primates, and rats (Ramus, Hauser, Miller, Morris, & Mehler 2000; Toro, Trobalon, & Sebastian-Galles, 2005). In the current study, we used the inspiration of reducing the phoneme set, but chose not to use the artificial pitch resynthesis. This is because the subtle prosodic properties that signal disfluency might not be preserved by the resynthesis process.

As with Experiment 1, if infants discriminate the fluent and disfluent samples, they should show a difference in their listening preferences for the two sample types. However, because we have manipulated the phonological properties of the samples, there is no longer a clear prediction regarding the direction of that preference. This

manipulation had two main effects: the words in the stimuli were no longer interpretable as English words, and the overall phonological complexity was decreased. Studies in which real English speech is used (like Experiment 1) have generally found familiarity preferences, or preferences for the more well-formed speech (e.g. Santelmann & Jusczyk, 1998; Shady, 1996; Soderstrom, 2002), whereas studies in which quasi-linguistic stimuli like nonsense syllables are used (as in Experiment 2) have often found novelty preferences, or preferences for the more ill-formed stimuli (e.g. Gomez, 2002; Gomez & Maye, 2005; Saffran et al., 1996). Regardless of the direction of preference, however, a significant difference in listening times will demonstrate that infants discriminate the samples.

Method

Participants.

The data included were from 24 infants between 20 and 23 months (range: 608-726, mean: 678 days, 13 male and 11 female), and 24 infants between 10 and 11 months (range: 300-344 days, 12 male and 12 female). Eight additional Post-verbal infants participated in the study but their data were discarded due to fussiness (2), lack of interest in the auditory stimuli (2), or average looking times less than 3 seconds (4). The data from 1 additional Pre-verbal infant were discarded due to fussiness, and the data from 4 additional Pre-verbal infants were lost due to a corrupted computer file. All participants were normally developing infants with normal hearing from Providence, RI, USA, and had parents and caregivers who were native speakers of American English.

Stimuli.

To remove lexical and grammatical information, the utterances from Experiment 1 were modified by reducing the phonemes to a set of manner changes. All fricatives became /s/, vowels became /a/, liquids became /l/, glides became /y/, stops became /t/, and nasals /n/. As noted, this method was based on the “saltanaj” transformation of Ramus & Mehler (1999). Thus, the example stimuli from (1) and (2), rewritten here as (3a) and (4a) became (3b) and (4b):

(3a) yeah , so we # ended up [///] we drove through Boston for like an hour, literally, with all the traffic +//.

(3b) ya, sa ya # antat at [///] ya tlas sla tastan sal lat an ayl, latalala, yas al sa tlasat +//.

(4a) yeah, so we ended up driving through Boston for like an hour, literally, with all the traffic.

(4b) ya, sa ya antat at tlasan sla tastan sal lat an ayl, latalala, ya al sa tlasat.

These stimuli were produced in a manner similar to that of Experiment 1 – the same trained speaker listened to the stimuli she had recorded for Experiment 1, and after careful practice produced the modified versions with prosodic properties as close as possible to the originals. These stimuli were then rated by a group of 11 adult listeners with the same rating scale as before (low-pass filtering was not used, as lexical and grammatical information was already removed from the samples). These raters were again able to discriminate the fluent from the disfluent stimuli with a high degree of accuracy ($t(10) = 7.2, p < .001$), confirming that the general prosodic differences between the fluent and disfluent passages had been preserved. The average rating for disfluent utterances was 5.2 and for fluent utterances 3.2. Only 2 of the 20 utterances had an average rating difference of less than .5. In order to preserve as much as possible the

prosodic cues to disfluency, it was decided not to alter the lengths of the audio files for this experiment, as the PSOLA method may degrade the quality of the audio. There was a less than 4% difference in average length across the stimuli. The average Pearson correlation of the pitch tracks between the normal and modified version was .58. Stimuli were concatenated into passages in a similar fashion to Experiment 1, with 1 s between each speech sample.

Procedure.

The procedure was identical to that of Experiment 1. Both pre- and Post-verbal infants were run using the video HPP apparatus.

Results and Discussion

Means and standard errors of looking times are shown in Figure 2. The older infants preferred the disfluent passages, looking to the disfluent passages for an average of 7.4 s ($SD = 3.39$) compared with 6.0 s ($SD = 2.09$) for the fluent passages. This was significant by 2-tailed paired t-test, $t(23) = 2.2$, $p < .05$, $d = .44$. Of the 24 infants in this condition, 11 infants preferred the disfluent passages on average by more than 1 s, while only 5 preferred the fluent passages by more than 1 s. The Pre-verbal infants again showed no preference for either passage type, looking to the disfluent passages for an average of 6.1 s ($SD = 2.22$) and to the fluent passages for an average of 5.9 s ($SD = 2.52$), $t(23) < 1$, $p > .5$, $d = .04$. Of the 24 infants in this condition, 5 infants preferred the disfluent passages on average by more than 1 s, while 5 preferred the fluent passages. A mixed ANOVA with age and passage type as factors found a main effect of passage type, $F(1, 46) = 4.15$, $p < .05$, and a marginal interaction between age and passage type, $F(1, 46) = 1.53$, $p = .088$.

Insert Figure 2 about here

In this second experiment, infants again discriminated between fluent and disfluent utterances. While the interaction between age and passage type was marginal, the individual t-tests suggest that like Experiment 1, Pre-verbal infants did not discriminate fluent from disfluent passages, but Post-verbal infants did discriminate. In contrast to the first experiment, however, the Post-verbal infants preferred the disfluent, rather than the fluent, passages. Reversals of preference have been found in other studies of this kind (Gomez & Maye, 2005; Thiessen et al., 2005) and are generally attributed to differences in age and/or complexity of the stimuli (Hunter & Ames, 1988). Based on these previous findings, it is likely that the reversal in preference in this case was due to the difference in phonological diversity in the two types of stimuli or to the type of linguistic processing each type of stimulus was likely to elicit. With natural language stimuli, infants were recognizing some familiar words and perhaps attempting to parse and comprehend the stimulus utterances. The “saltanaj” stimuli contained no familiar words and no basis for parsing or comprehension and were thus susceptible to analysis only at the levels of phonology and prosody. It is noteworthy that infants showed greater attention to the well-formed stimuli in case of the familiar, natural language stimuli, just the situation that would be most beneficial to the language learner. Although other explanations for the reversal in preference are possible, the relevant finding in this experiment is that Post-verbal infants do indeed discriminate fluent and disfluent speech samples in the absence of lexical and grammatical cues. This supports the idea that infants use the prosodic characteristics of speech to discriminate well-formed from ill-

formed speech, although other cues, such as word repetition, part-words and *ums* and *uhs* may also be used.

General Discussion

Across two experiments, we have shown that Post-verbal, but not Pre-verbal, infants discriminate between fluent and disfluent adult-directed speech, and can do so even when lexical and grammatical information are removed from the stimuli.

Furthermore, when this speech is in English, infants show a preference for the fluent speech. These results are important because of the strikingly high correlation between ungrammaticality and disfluency in adult-directed speech: if infants filter out disfluent adult-directed speech, they may thereby exclude virtually all ungrammaticality from their input.

Our findings suggest a number of things about the way that infants process speech. First, Post-verbal infants clearly attend to adult-directed speech, at least in some contexts. If not, they could not have succeeded at our task, since only one of the 20 utterance pairs was infant-directed. Second, the ability to detect disfluency may develop as infants mature linguistically. Third, the *prosodic* properties of disfluencies influenced older infants' listening preference, since they discriminated the samples even when other cues were removed.

Our finding that infants detect disfluency in AD speech (and are therefore able to process some properties of AD speech) highlights the need for further research into the role that AD speech may play in early language development. Given that the majority of speech spoken in an infant's presence may be directed toward older children or other

adults (van de Weijer, 2002), this input may play an important role in the linguistic development of infants. The preference that infants, at least younger infants, show for characteristically ID speech does not demonstrate that infants are completely insensitive or indifferent to the presence of AD speech. In fact, recent evidence suggests that infants as young as 18 months may learn novel words from speech directed from adult to another (Akhtar, 2005; Floor & Akhtar, 2006). Furthermore, examining the interaction that infants in our culture may have with adult-directed speech may allow us to begin to address the puzzles of language acquisition in societies where ID speech as a social behavior may not be the norm.

While the individual age-by-passage type interactions in our study only approached statistical significance, there was a clear difference in the pattern of significant preferences found in the older, linguistically mature age group, and the non-significant looking time differences found with the younger, preverbal age group. These differences suggest that sensitivity to disfluency, even at the prosodic level, may develop gradually. Given that previous studies have found that infants much younger than 10 months are highly sensitive to general prosodic characteristics of fluent speech (e.g. Mandel et al., 1994, 1996; Nazzi et al., 2000; Ramus et al., 2000; Soderstrom et al., 2003) there are a number of possible explanations for the lack of discrimination by the preverbal age group in the current study. One possibility is that the younger infants were less interested, and therefore attended less well to the adult-directed stimuli than the older infants. However, infants in the two age groups had similar overall listening times to the passages, and the older infants had, if anything, a *higher* rate of fussiness than the younger age group, neither of which supports an explanation based on interest alone.

Attentional differences may still have played a role in terms of the particular aspects of the speech signal to which the infants of different ages were sensitive. Alternatively, discovering the particular prosodic cues that signal disfluency may be something that is learned, and younger infants have simply not acquired the linguistic knowledge to discover these cues yet.

If this latter possibility holds, it may be that the linguistic maturity of the older infants not only creates a need for the more complex input provided by AD speech, but is also a prerequisite for developing the ability to differentiate the prosodic characteristics of disfluency. It may be that infants may first begin to detect disfluencies by noting speech input that violates grammatical and lexical expectations in relatively simple sentences and only later learn the prosodic characteristics associated with these disfluencies. If so, there should be an intermediate age at which infants are able to detect disfluencies in simple sentences, but not if the lexical and grammatical cues are removed.

The manner in which infants compensate for disfluency in the input may also be something that develops. Adults, when confronted with disfluency, not only detect the error but may also implicitly make the necessary repairs to understand what was said. It is for this reason that the high rate of disfluency in every day speech is difficult to notice consciously unless it is specifically pointed out. In the current study, we have shown that infants detect these disfluencies, but not what they do with the information. Early on, infants may simply discard ill-formed utterances in favor of well-formed utterances. Ultimately, however, mature language users must learn to make the online repairs necessary for fluent listening. How and when this occurs is an open question.

Chomsky's and others' concerns about the nature of the input (e.g. Baker, 1979; Chomsky 1962, 1965; Peters, 1972), while not totally unfounded, may be overstated. The speech environment to which infants are exposed in daily life may not resemble the articulate, well-formed utterances found in radio addresses and laboratory experiments. But both adult speakers and infant listeners may contribute to minimizing the impact of the chaos of disfluencies and everyday speech errors. Adult speakers, at least in many language communities, use a specialized speech register when speaking with infants that minimizes errors and provides simple, well-formed input. When speakers do make speech errors, prosodic information helps to indicate their presence and minimize the damage they might inflict on the listener, whether a mature listener, or a language learner.

Infants, for their part, may be viewed as discerning listeners. Rather than absorbing all speech acts indiscriminately, they may selectively attend to speech input that is most helpful or a well-formed representation of the underlying grammar. Given the previous research showing a preference for ID speech, young infants may attend exclusively (or at least primarily) to the short, well-formed utterances characterized by ID speech. Older infants may expand the scope of what they process to include more AD speech, but limit themselves to well-formed, fluent-sounding utterances. Our findings support this view by demonstrating that these older infants are sensitive to the prosodic characteristics that differentiate fluent from disfluent speech. Using such attentional strategies, infants could avail themselves of the best the input has to offer, and ignore the worst.

Author Note:

The authors wish to thank the parents and their infants for participating in the study. They also wish to thank Katherine White and two anonymous reviewers for comments on a previous version of the manuscript – mistakes remain our own. This research was funded by NIH grants 1RO1HD32005 awarded to JLM and 5F32HD042927 to MS.

Table 1: Fluent and disfluent utterance-internal prosodic breaks in the transcript

	# breaks-total	# breaks-AD	# breaks-ID
Fluent	273	151	122
Disfluent	96	93	3

Table 2: Number of grammatical and ungrammatical phrases preceding fluent and disfluent prosodic boundaries

	Fluent	Disfluent
Grammatical	474	19
Ungrammatical	6	80

Table 3: Average properties of disfluent and fluent utterances

	Duration Pre-adjustment (ms)	Syllables	Repeated Words	Prosodic boundaries
Fluent	3565	17.6	0.85	1.85
Disfluent	3601	17.25	0.95	2.8
Significance	n.s.	n.s.	n.s.	p < .001

Figure 1: Mean looking times to disfluent and fluent speech in Experiment 1

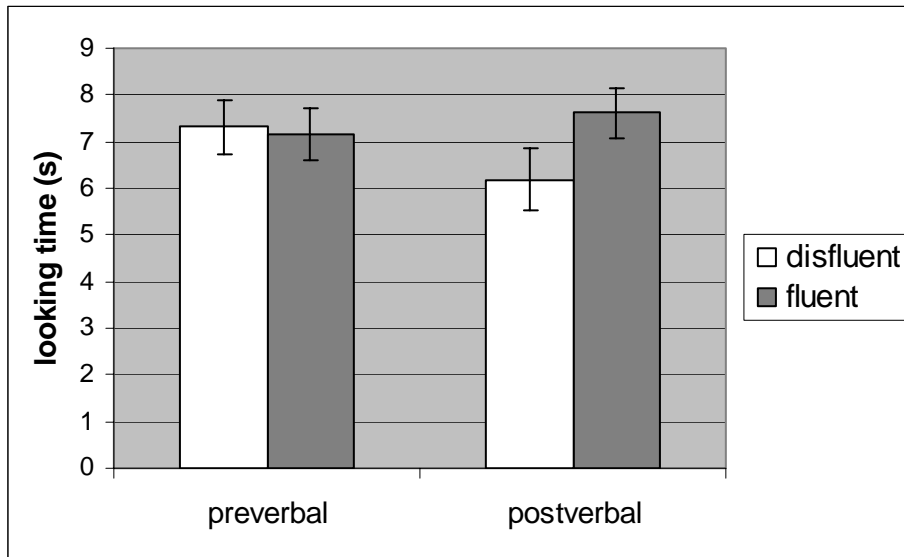
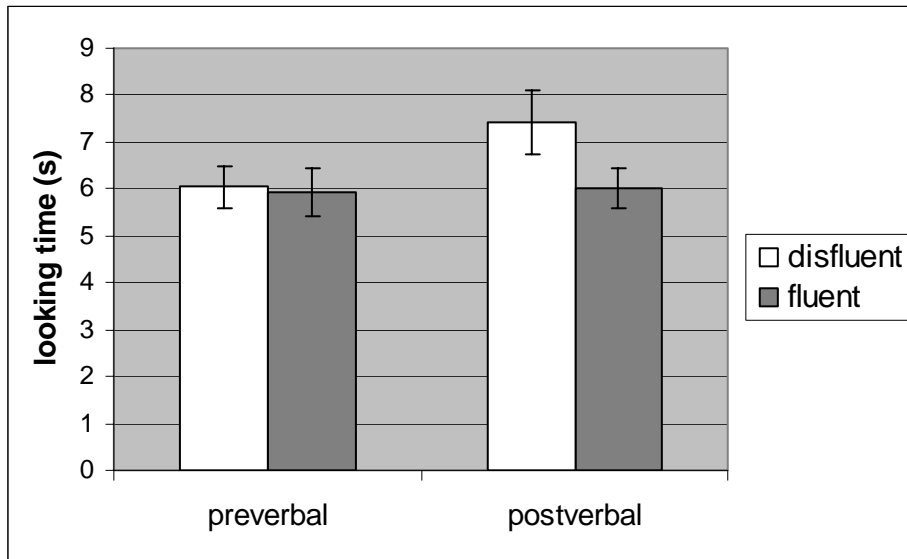


Figure 2: Mean looking times to disfluent and fluent speech in Experiment 2



Appendix

D: Cause sometimes I put a little bow in her &hea [/] hair and people still think she's a boy.

F: Cause sometimes I put a little bow in her hair, and people still think she's a boy.

D: But Magnolia [///] he [///] my husband jokes, he calls it South Dallas.

F: But my husband jokes, he calls Magnolia South Dallas.

D: Well # um # they uh [///] # no we were able to switch no problem.

F: Well, no. we were able to switch, no problem.

D: So, yeah we were very lucky , it's just [///] we've just been waiting here.

F: So yeah, we were very lucky. It's just that we've been waiting here.

D: They opened up the roads at six a.m. &an [///] but by then it was too late, we had already +//.

F: They opened up the roads at six a.m., but by then it was too late, we had already left.

D: Oh , &traf [///] and then +//.

F: Oh, traffic, and then more.

D: You &eh [//] would you grab that pacifier, babe?

F: Would you grab that pacifier, babe?

D: So we were like, well fat lot of good it does running all the # you know, uh, freeways under the city +//.

F: So we were like, well fat lot of good it does running all the freeways under the city, you know.

D: Yeah, so we # ended up [///] we drove through Boston for like an hour, literally, with all the traffic +//.

F: Yeah, so we ended up driving through Boston for like an hour, literally, with all the traffic.

D: We couldn't figure out &n [/] none of the &sign [//] you know, the old tunnel +...

F: We couldn't figure out any of the signs, you know, the old tunnel.

D: And it &wa [//] there was traffic # down there, like cause they were [//] they routed everybody through the city and so there was traffic, even at +//.

F: And it was crazy. There was traffic down there, like cause they were routing everybody through the city and so there was traffic even at night.

D: Yeah, we're going to Magnolia too, which I said [///] # that's +//.

F: Yeah, we're going to Magnolia too, which I said. That's right.

D: That's the &uh [//] that's the one that's next to them, but +//.

F: *That's the one. That's the one that's next to them, but further.*

D: You know and that's what &I [//] we just [//] I was like Magnolia used to be so far, nobody lived in Magnolia and now +//.

F: *You know and that's what I knew. I was like Magnolia used to be so far, nobody lived in Magnolia and now everyone does.*

D: That's one of those things well # we've got this pillow for her which is really nice but # uh, oh that really [//] cause I'm one of those [//] I always grab a pillow, I sleep when I +//.

F: *That's one of those things. Well we've got this pillow for her which is really nice, but that's really hard, cause I'm one of those people, I always grab a pillow, I sleep when I fly.*

D: I've traveled on, yeah, that before and &no [//] you can't with her anymore.

F: *I've traveled on that before, yeah, and no, you can't with her anymore.*

D: And she did [/] she did really well, she [//] once she +//.

F: *And she did well. She did really well, once she slept.*

D: The ten [/] the ten thirty flight or +//.

F: *The ten or the ten thirty flight.*

D: I think [/] yeah.

F: *I think so, yeah.*

D: Is that a [//] that's another traveling baby. (ID speech)

F: *Is that a baby? That's another traveling baby. (ID speech)*

Key:

#: pause

[/], [//], [///]: retracing without correction, retracing with correction, reformulation

+/. , +//. , +...: interrupted utterance, self-interruption, trailing off

&: word fragment

References

- Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, 8(2), 199-209.
- Aylett, M. P. (2005). *Extracting the acoustic features of interruption points using non-lexical prosodic analysis*. Paper presented at the Disfluency in Spontaneous Speech Workshop (DiSS '05), Aix-en-Provence, France.
- Baker, C. L. (1979) Syntactic theory and the projection problem. *Linguistic Inquiry*, 10, 533-582.
- Bartkova, K. (2005). *Prosodic cues of spontaneous speech in French*. Paper presented at the Disfluency in spontaneous speech workshop (DiSS '05), Aix-en-Provence.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341-345.
- Broen, P. A. (1972). *The verbal environment of the language-learning child*. Washington, DC: American Speech and Hearing Association.
- Chomsky, N. (1962). Explanatory Models in Linguistics. In E. Nagel, P. Suppes & A. Tarski (Eds.), *Logic, Methodology and Philosophy of Science* (pp. 528-550). Stanford, CA: Stanford University Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73-111.
- Cole, J., Hasegawa-Johnson, M., Shih, C., Kim, H., Lee, E.-K., Lu, H.-y., et al. (2005). *Prosodic parallelism as a cue to repetition and error correction disfluency*. Paper

- presented at the Disfluency in Spontaneous Speech Workshop (DiSS '05), Aix-en-Provence, France.
- Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, 20(4), 477-488.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584-1595.
- Elman, J. L. (2005). Connectionist models of cognitive development: where next? *Trends in Cognitive Sciences*, 9(3), 111-117.
- Fernald, A. (1985). Four-month-olds prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-293.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & 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.
- Floor, P., & Akhtar, N. (2006). Can 18-month-olds learn words by listening in on conversations? *Infancy*, 9(3), 327-339.
- Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge, UK: Cambridge University Press.
- Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431-436.

- Gomez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2), 183-206.
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research*, 44, 1189-1200.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The Origins of Grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.
- 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.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.
- 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.
- Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The Headturn Preference Procedure for testing auditory perception. *Infant Behavior and Development*, 18, 111-116.
- Levelt, W., & Cutler, A. (1983). Prosodic marking in speech repair. *Journal of Semantics*, 2, 205-217.
- Lewis, J. D., & Elman, J. L. (2001). Learnability and the Statistical Structure of Language: Poverty of the stimulus arguments revisited. In B. Skarabela, S. Fish &

- A. H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (Vol. 1, pp. 359-370). Somerville, MA: Cascadilla Press.
- Lickley, R. J., & Bard, E. G. (1998). When can listeners detect disfluency in spontaneous speech? *Language and Speech*, *41*(2), 203-226.
- Loevgren, T., & van Doorn, J. (2005). *Influence of manipulation of short silent pause duration on speech fluency*. Paper presented at the Disfluency in Spontaneous Speech Workshop (DiSS '05), Aix-en-Provence, France.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (third ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants to organize and remember speech information? *Cognition*, *53*, 155-180.
- Mandel, D. R., Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development*, *11*, 181-196.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143-178.
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*, 67-85.

- Nakatani, C. H., & Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *Journal of the acoustical society of america*, 95(3), 1603-1616.
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, 1, 123-147.
- Newman, R. S. (2006). Changes in preference for infant-directed speech in low and moderate noise by 5- to 13-month-olds. *Infancy*, 10(1), 61-76.
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 109-149). Cambridge, UK: Cambridge University Press.
- 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 and Development*, 15, 325-345.
- Peters, S. (1972). The projection problem: How is a grammar to be selected? In S. Peters. (Ed.) *Goals of linguistic theory* (pp. 171-188). Englewood Cliffs, NJ: Prentice Hall.
- Ramus, F., Hauser, M., Miller, C., Morris, D. & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349-351.

- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the acoustical society of america*, *105*(1), 512-521.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*, 105-134.
- Schieffelin, B. B. (1985). The acquisition of Kaluli. In D.I.Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1: The data, pp. 525-593). Hillsdale, N.J.: Erlbaum.
- Shady, M. (1996). *Infants' sensitivity to function morphemes*. Unpublished dissertation, State University of New York, Buffalo, NY.
- Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. A. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English. *Journal of Speech, Language, and Hearing Research*, *41*, 874-886.
- Soderstrom, M. (2002). *The acquisition of inflection morphology in early perceptual knowledge of syntax*. Unpublished dissertation, Johns Hopkins University, Baltimore, MD.
- Soderstrom, M., Blossom, M., Morgan, J. L., & Foygel, I. (under revision). The speech environment of 6-10 month old infants. *Journal of Child Language*.
- Soderstrom, M., & Morgan, J. L. (2005). *Disfluency in speech input to infants? The interaction of mother and child to create error-free speech input for language*

- acquisition*. Paper presented at the Disfluency in Spontaneous Speech Conference (DiSS '05), Aix-en-Provence, France.
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49*, 249-267.
- Soderstrom, M., White, K. S., & Conwell, E. (2005). Evidence for grammatical knowledge of content words in 16-month-olds, *30th Annual Boston University Conference on Language Development*. Boston, MA.
- Spence, M. J., & Freeman, M. S. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development, 19*, 199-212.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*(1), 53-71.
- Toro, J. M., Trobalon, J.B., Sebastian-Galles, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 95-100.
- van de Weijer, J. (2002). *How much does an infant hear in a day?* Paper presented at the GALA 2001 Conference on Language Acquisition, Lisboa.
- 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*(2), 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.