Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop.
10–12 September 2005, Aix-en-Provence, France, pp. 157-162.

Disfluency in speech input to infants? The interaction of mother and child to create error-free speech input for language acquisition

Melanie Soderstrom[*] *& James L. Morgan*[*]

[*] Brown University, Providence, USA

Abstract

One characteristic of infant-directed speech is that it is highly fluent compared with adult-directed speech. However, the speech that infants hear still contains disfluencies. Such disfluencies might potentially cause problems for infants during language development. We first analyzed samples of spontaneous speech in the presence of infants (both adult- and infant-directed) and found that under ideal circumstances the speech infants hear is highly fluent. Under less than ideal circumstances infants hear much more highly disfluent speech - however this disfluent speech is almost entirely adult-directed. While grammatically ill-formed, the prosodic structure of these disfluencies might signal their ill-formedness to the infants. In a preference experiment, 10 month olds listened longer to infant-directed speech samples containing prosodic disfluencies than to equated samples without disfluency. However, this effect was found in only one of two counterbalancing groups. Using adult ratings of low-pass versions of these speech samples, we found that infants' preferences were correlated with the adults' perception of the relative disfluency of the samples. A follow-up experiment using adult-directed disfluencies found that while the 10 month olds showed no differences in their listening preferences, older infants preferred to listen to the fluent speech. These results suggest that younger and older infants attend differently to infant and adult-directed speech, and that older infants may be able to differentiate grammatical adult-directed input from input distorted by disfluency. We discuss implications of these findings for language acquisition.

1. Introduction

The extent to which the disfluent character of speech might pose a problem for the learner has been a subject of great interest in the study of language acquisition. Chomsky [1] famously asserted that the language input to the child included many "interrupted fragments, false starts, lapses, slurring, and other phenomena that can only be understood as distortions of the underlying pattern". While typical adult-adult speech might well fit this description, studies of child-directed speech have found it to be highly fluent. For example, Newport, Gleitman & Gleitman [8] found only 1 disfluent child-directed utterance out of 1500, and only 4% of utterances were untranscribable due to mumbling or slurring. Even the adult directed speech in this study was fairly fluent - 5% disfluent and 9% untranscribable. This finding, which has been supported by numerous studies of child-directed speech, suggests that the input to the child might not be subject to the distortions that characterize normal adult-directed speech.

Nevertheless, the input to the language learner may not be unswervingly well-formed. For one thing, a recent estimate of the total language input to an infant found that only 15% of the speech heard by the infant is directed to that infant - an additional 30% was directed at an older sibling [12]. Therefore, if the ambient adult speech is processed by infants, between 55 and 85% of the language input in the earlier stages of grammatical development might be of the type described by Chomsky.

Even if infants do not process a significant number of explicitly ungrammatical utterances due to disfluency, the ill-formed prosodic structure of disfluent utterances may be a source of misinformation about the grammatical structure of the language. A growing body of literature suggests that infants as young as 2 months are highly sensitive to the prosodic structure of utterances [e.g. 5, 6] and by 6 months can use prosodic information to organize fluent speech [e.g. 7, 9, 10] into grammatically-relevant units, even before lexical information (including word boundaries) is available. This "prosodic bootstrapping" theory depends on the *prosodic* well-formedness of the input to the language learner. Studies such as [8] examined only major errors in grammatical structure of the utterance, what they referred to as "true garble", but not disfluencies in the phonological or acoustical structure of the utterance. Such prosodic disfluencies might themselves be a critical source of error in this initial process of grammatical development. On the other hand, if infants can detect these prosodic disfluencies, they may provide cues to the infant about the reliability of the utterances as a source of information about the grammar.

The current study asks three questions about the possible effect of disfluency on early stages of language acquisition:

(1) Does infant-directed speech contain prosodic disfluencies?

(2) Are infants able to differentiate between fluent prosodic breaks and prosodic disfluencies in infant- and adult-directed speech?

(3) Are there developmental differences in infants' sensitivities?

In the second and third parts of this paper, we will provide some analyses of the presence and properties of disfluencies in the speech input of infants in different environments. In the fourth and fifth parts, we will examine behaviorally adults' and infants' sensitivity to the prosodic properties of disfluencies in speech.

2. Corpus analysis I: Best case scenario

2.1. The corpus

Maternal speech to two infants was collected every 1-3 weeks in a natural home setting while the infant was 6-10 months old. As part of a larger study [11], both mothers were asked to record about an hour per week of speech, and to make recordings at least 30 minutes long, however some recordings were shorter than this. In total, approximately 8.5 hours were collected during this period for the first mother (MOT1), and 14 hours for the second mother (MOT2). These recordings were then transcribed using the CHAT transcription system [4], for a total of 9067 utterances for MOT1 and 10604 utterances for MOT2. The majority of these utterances were infant-directed. However, for MOT1, there were also a significant number of adult- and child-directed utterances. These are reported separately.

Each utterance was then coded for a variety of syntactic and prosodic features. In particular, major utterance-internal prosodic breaks (i.e., anything transcribed with a comma, #, [/], etc.) were coded qualitatively as follows:

(1) F: Any fluent, prosodic break at the conjunction of two well-formed grammatical units.

(2) D: Any disfluent prosodic break caused by restart, reformulation, speech error, etc.

(3) O: Any prosodic break not caused by restart, reformulation, etc., that was otherwise odd or ill-formed (primarily pauses for thought). In this first analysis this category included prosodic breaks which were odd either because they were prosodically ill-formed, or because they occurred at a grammatically inappropriate location.

2.2. Prosodic disfluencies in speech samples

Our samples of speech to two young infants found relatively few examples of utterance-internal prosodic disfluencies compared with utterance-internal fluent prosodic breaks (Table 1). These data suggest that only about 5-10% of prosodic breaks are disfluent in infant-directed speech (Formula: # D + O breaks / total # prosodic breaks).

Table 1: Fluent, disfluent, odd, and total prosodic breaks in infant-directed speech (per utterance). Bottom Row: Disfluent breaks per total breaks).

	MOT1	MOT2	MOT1-IDonly
F/utt	.390	.219	.330
D/utt	.013	.008	.007
O/utt	.013	.014	.007
Tot/utt	.416	.241	.344
(D+ O)/tot	.063	.091	.041

These data present a relatively good picture for prosody as a source of information for infants about the structure of their language. However, these data provide in some sense a "best case scenario" picture of the speech input environment of infants. These recordings were taken at home, in relative quiet (although for some parts of the MOT1 recordings, the father and siblings were present and provided a noisier input environment). Not all of an infant's speech input will be obtained in such ideal conditions. For example, INF1 (MOT1's infant) was in a busy daycare during the day, and was exposed to a variety of children and caregivers. While we do not have recordings of this environment, we do have a recording by another mother in the larger study, MOT4, during a wait in a busy airport terminal. This transcript allows us to examine a more complex input environment.

3. Corpus Analysis II: Worst Case Scenario

3.1. Transcript

This transcript was obtained while the infant was 2.5 months old. Mother, father and infant were waiting in an airport terminal for their flight to be called. The mother chose this time to make one of her recordings. During this time a variety of other people in the terminal interacted with the mother, so a smaller percentage of the utterances in this recording were directed at the infant. This environment was likely to generate a much larger number of disfluencies not only because many of the utterances were adult-directed, but also because of the high ambient noise level, and because the mother was engaged in an animated discussion with strangers.

Due to the greater number of disfluencies overall, we distinguished in this analysis between prosodically and syntactically odd breaks. A prosodic break was only considered an O if it was *prosodically* odd. If it sounded fluent but was grammatically inappropriate, it was classified as a C. Utterances were then separately analyzed for the grammaticality of the word groupings caused by both fluent and disfluent prosodic breaks.

(1) C: Any fluent-sounding prosodic break, regardless of the grammatical context.

(2) R: Any disfluent prosodic break caused by restart, reformulation, speech error, etc.

(3) O: Any prosodic break not caused by restart, reformulation, etc., that was prosodically odd or ill-formed (primarily pauses for thought).

3.2. Prosodic disfluencies in a complex environment

Clearly, this mother produced both a larger rate of prosodic breaks per utterance overall, and also a much larger proportion of disfluent breaks (Table 2). However, the overall rate of breaks she produced in infant-directed speech was comparable to that of the other two mothers, and if anything, the rate of infant-directed disfluent breaks was smaller.

Table 2: Absolute numbers and prosodic breaks per utterance for MOT4

	# breaks	breaks/utt	# breaks-ID	breaks/utt-ID
C	273	.503	122	.396
R	42	.077	2	.006
O	54	.099	1	.003

Proportionally, this mother's infant-directed disfluencies accounted for only 2% of the total infant-directed prosodic breaks in the speech samples. However, looking at the total sample, this proportion is much higher, up to 26%.

Table 3 gives the percentages of phrases and clauses that are grammatically ill-formed when we consider utterances, all prosodically bound word sequences (i.e. word sequences to either side of a C, R, or O), only phrases and clauses bound on at least one side by an R or O break, and finally phrases and clauses bounded only by well-formed prosodic breaks. Isolated single words were excluded from this and the following analysis, because they are by defnition well-formed grammatical units.

Table 3: Percentages of syntactically ill-formed units by prosodic context for MOT4

	Utterances	All pros units	R's	O's	C only
Phrases	.054	.228	.778	.375	.028
Clauses	.157	.151	.467	.640	.080

Clearly, the mother in this transcript is producing a large number of syntactically ill-formed prosodic units due to disfluency. However, the vast majority of prosodically well-formed word sequences (the "C only" column) are also syntactically well-formed (97% of phrases and 92% of clauses). These data suggest that the ability to detect prosodic disfluency might be of great benefit to infants in determining which word sequences constitute reliable input for language acquisition.

The picture looks even better if we examine the relative number of grammatically well-formed and ill-formed phrases/clauses preceding fluent and disfluent prosodic breaks, including disfluent utterance endings (which were not counted in Tables 2 and 3). Ninety-nine percent of word sequences preceding fluent prosodic breaks are grammatically well-formed. Of the 6 ungrammatical units preceding a fluent-sounding prosodic break, 4 are directly following a disfluency. Of the remaining two, one was formed by the insertion of ”you know” in the middle of a phrase. The other was preceded by an ”um”, indicating disfluency, but simply did not sound disfluent to the ear.

Table 4: Number of grammatical and ungrammatical phrases preceding fluent and disfluent prosodic boundaries.

	Fluent (C)	Disfluent (R)	Odd (O)	Disfluent Endings
Gramm.	474	4	8	7
Ungramm.	6	29	27	24

Based on our transcribers’ judgments, MOT4’s adult-directed speech in this sample, though highly disfluent, also contains highly reliable cues to grammaticality. The following section examines whether these prosodic cues to disfluency are detectable by naive adult listeners, and more importantly, by infants.

4. Differentiating fluent and disfluent speech: Infant-directed speech

The data from the preceding section suggest that in some circumstances, infants might be exposed to a large amount of disfluency in speech. However, the ungrammatical word groupings cause by this disfluency appeared to be reliably marked prosodically. We examined both adults’ and infants’ ability to detect prosodic cues to disfluency.

4.1. Stimuli

Disfluent utterances directed at the infant or older siblings were culled from the transcripts of MOT1. Utterances were discarded if they contained background noise, were not clear enough to be fully transcribed, or if they were highly repetitive due to restarting. Due to these selection limitations and the small number of disfluencies overall, we were not able to be particular about our disfluent utterance types. These included a variety of utterance types, including questions and declaratives, and utterance lengths. They covered a range of disfluency types, including interruptions, repetitions, restarts and reformulations.

Because these utterances were created in a spontaneous speech environment, it was not possible to obtain paired fluent samples which matched the disfluent samples on variables such as length, intensity, number of syllables, prosodic structure etc. A pilot attempt to directly use the speech samples from the recordings as fluent controls was deemed too variable. Therefore, in order to control for these other factors, both fluent and disfluent utterances used in the experiment were created in the laboratory, and produced by the first author. Disfluent utterances were produced by listening to, and mimicking as closely as possible, the disfluent speech of the mother. Fluent utterances were created by transforming the disfluencies into fluent prosodic breaks.

For example, the fluent utterance (2) was created from the disfluent utterance fragment (1):

(1) Should we start +//. Oh, a big yawn from you.
(2) Should we start? Oh, a big yawn from you.

Utterances were grouped into passages of 6 utterances each. Eight disfluent passages and eight matched fluent passages were produced. Each passage was about 20 seconds in length. While mimicked disfluencies are probably not prosodically identical to real disfluencies, we felt that this method was the best compromise between a controlled experiment and using stimuli that are most representative of the kinds of disfluencies that infants are likely to hear. It is likely, however, that this design underestimated the prosodic differences between fluent and disfluent speech - this will be pursued further below.

Table 5 provides information about the properties of these infant-directed disfluent utterances and their fluent matched controls.

Table 5: Average properties of infant-directed disfluent and fluent utterances

	Length (ms)	Syllables	Repeated words	Prosodic boundaries
Fluent	2438	12.0	0.313	1.23
Disfluent	2459	12.1	0.354	1.61

There were no significant differences between length, number of syllables, or number of repeated words. The disfluent passages contained significantly more utterance-internal prosodic boundaries than the fluent passages ($t(47)=4.08$, $p < .001$). Along with prosodic disruptions, the disfluent passages contained 1 um/uh and 18 part-words.

4.2. Adult Ratings

In order to measure how disfluent our “disfluent” samples were, we presented the samples to 8 adult raters, who were asked to judge how “fluent” or “disfluent” they sounded. This served two purposes - to establish that our disfluent samples did indeed contain prosodic cues to their disfluency, and to ascertain that our “fluent” controls actually sounded more fluent.

Samples were low-pass filtered at 400 Hz (with 100 Hz smoothing) so that the raters heard only the overall prosodic character of the speech samples, and not the high-order phonological or lexical information. Raters were asked to rate each individual utterance from the speech samples, from 1 (highly fluent) to 4 (highly disfluent). All 8 raters scored the disfluent utterances on average as more disfluent than the paired fluent utterances. This difference was highly significant by two-tailed t-test ($t(7)=7.688$, $p < .001$), suggesting that the disfluent and fluent utterances were discriminable based on prosodic characteristics. However, the difference in rating was not very large - an average of 2.63 for disfluent utterances, and 2.40 for fluent utterances. In order to try to pull out this difference, a second set of ratings was obtained with a larger (1-7) scale, but the means were similarly close (4.8 versus 4.1 on the wider scale), so the original ratings were used in subsequent analyses.

This result not only validates our stimuli, but also provides evidence that disfluencies in infant- and child-directed, while rare, are detectable based on their prosodic characteristics alone - at least by adult listeners. We next examined whether infants are sensitive to these prosodic cues.

4.3. Participants

Thirty-two 10 month old infants participated in the experiment. They ranged in age from 309 days to 342 days. There were fifteen males and seventeen females. An additional 2 infants were tested but not included in the experiment due to fussiness. 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.

4.4. Design

If infants are able to differentiate fluent and disfluent speech samples, they are likely to exhibit differences in their listening preferences between the two passage types. Because it is often difficult in behavioral research to predict whether infants will show a preference for the ill-formed stimulus (a novelty effect) or the well-formed stimulus (a familiarity effect), a two-tailed statistical test is employed to determine whether infants show a difference in their preferences, regardless of the direction of that difference.

Infants were divided into two counterbalancing groups. Prior to the test phase there was a pre-test phase to familiarize the infants with the procedure. This consisted of two trials of speech stimuli, one of which was a 10 second repetition of a fluent, prosodically well-formed phrase; the other was a 10 second repetition of a prosodically ill-formed phrase-like word sequence. The stimuli came from a separate study [9], and were chosen to minimize the impact of the pre-test phase on infants' preferences during the test phase. Each group then heard all 8 test passages. Group 1 heard the disfluent versions of passages 1-4 and the fluent versions of passages 5-8, while group 2 heard the fluent versions of passages 1-4 and the disfluent versions of passage 5-8. The passages were presented once each in a random trial order.

4.5. Procedure

Infants were tested using the Headturn Preference Procedure. In this procedure, a flashing light is paired with a sound stimulus. When the infant looks toward the flashing light, the trial is initiated and the sound stimulus begins to play. When the infant looks away for at least 2 consecutive seconds, the trial ends. Infants' preference for the difference stimulus types is measured by their total orientation time toward the flashing light on each trial. For further details on this method, see [2].

4.6. Results

Across the two groups, infants preferred the disfluent passages, but this difference was non-significant. However, an examination of the two groups found that while group 1 showed no difference in their preferences, group 2 significantly preferred the disfluent passages (8.9 s mean looking time) to the fluent passages (7.4 s) ($t(15) = 2.174$, $p < .05$, two-tailed). This result is suggestive that infants do differentiate fluent and disfluent speech. But given the lack of statistical significance across the two groups, they must be interpreted with caution. However, recall that the way the stimuli were created may have minimized the prosodic differences across the sample types. If so, it is not surprising that infants would show only a weak difference in their preferences. Furthermore, because of the stringent criteria applied to the samples taken from the transcripts, there was some difficulty finding enough disfluent utterances to use. The best utterances were selected from a larger pool of possible samples. Therefore, the later-chosen samples were potentially less disfluent overall, and these were the samples that group 1, which showed no differences, heard as disfluent.

Because only one group of infants had a listening preference, we wished to confirm that the infants' preference was based on the fluent/disfluent distinction we were investigating. We therefore performed a correlational analysis between the infant listening preferences and the adult ratings. Because the infant listening preferences were obtained for the sample passages and the adult ratings on the individual utterances, we took an average of each rater's scores across the utterances in the sample passages. These 16 averaged scores were used in a Pearson correlation with the infant listening preferences. There was a weak but significant correlation ($r = .441$, $p < .05$, one-tailed) overall. Looking by group, Group 2's listening preferences were significantly correlated with the ratings ($r = .749$, $p = .016$, one-tailed), while Group 1's listening preferences were not correlated ($r = .031$, $p > .25$).

5. Differentiating fluent and disfluent speech: Adult-directed speech

The infant behavioral results with infant-directed speech were equivocal. While there was a significant difference in listening preferences for one of the two groups, the effect was not significant across the two groups. Might the infants show a more reliable listening preference with adult-directed stimuli? On the one hand, the disfluencies in the adult-directed speech may be more salient as well as simply more numerous. On the other hand, because these utterances are adult-directed, the speech is more rapid, and the prosodic characteristics of the speech are less salient. Furthermore, infants show greater attention to infant-directed than adult-directed speech overall [2]. Therefore, the disfluencies in adult-directed speech may be more difficult to detect. If infants show a reliable difference in their preferences for the disfluent and fluent versions of these adult-directed utterances, it will add support to our claim that infants differentiate disfluent from fluent speech, particularly in the speech mode for which it would be most beneficial because of the greater presence of disfluencies - i.e., adult-directed speech. Furthermore, if infants show a more reliable difference with the adult-directed samples, this would be a unique case where the properties of adult-directed speech might actually be more beneficial to infants than those of infant-directed speech.

5.1. Stimuli

Forty-two disfluent utterances were culled from the same transcript used in the "worst case scenario" analysis. Criteria for selection, and the process for creating the final stimuli were similar to those for infant-directed speech. However, the constraint on background noise in the sample was relaxed as long as the utterance was fully intelligible. We also included utterances with more repetition in order to maximize the disfluency in the samples.

5.2. Adult Ratings

We used the same procedure as with the infant-directed stimuli to obtain ratings from 8 adult listeners, using the 7 point scale. As with the infant-directed stimuli, 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 4.52 and for fluent utterances 3.26. Because we had a larger selection of utterances, we then chose the utterances with the largest differences as the stimuli for the behavioral study. Five utterances were selected for each of four fluent and disfluent passage pairs. All utterances selected had a difference of at least 1 point between their average fluent and average disfluent rating. Thus, in this study, the difference between the fluent and disfluent versions was more extreme in terms of adults' perceptions. Also, in order to better control for the small differences in length between the fluent and disfluent versions, the versions were equated for length by slowing down the shorter of the two and speeding up the longer of the two. These doctored stimuli were then rated a second time to verify that they were still highly discriminable by fluency.

Table 6 provides information about the properties of these adult-directed disfluent utterances and their fluent matched controls.

Table 6: Average properties of adult-directed disfluent and fluent utterances (length is pre-adjustment)

	Length (ms)	Syllables	Repeated Words	Prosodic boundaries
Fluent	3565	17.6	0.85	1.85
Disfluent	3601	17.25	0.95	2.8

As with the infant-directed utterances, there were no significant differences between starting length, number of syllables, or number of repeated words. The disfluent passages contained 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.

5.3. Participants

Twenty-three 10 month old infants participated in the experiment. They ranged in age from 309 days to 339 days. There were eleven males and twelve females. Fourteen older infants, ranging from 20 to 24 months, also participated, in an ongoing experiment. There were six males and eight females in this age group. An additional 4 infants at this age were tested but not included in the experiment due to fussiness. 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.

5.4. Design & Procedure

The design and procedure were identical to the first experiment for the 10 month olds, except that the fluent and disfluent version of each passage were both presented to each infant. So each infant heard all four disfluent passages, and the corresponding four fluent passages. For the 22 month olds only, a variant of the testing procedure was used in which an image on two T.V. screens are paired with the sound stimuli instead of flashing lights. This was done to maintain the interest of the older infants throughout the testing session.

5.5. Results

The 10 month olds showed no listening preferences for either the disfluent or fluent stimuli. Mean looking time to the disfluent passages was 7.3 s and to the fluent passages 7.1 s ($p > .5$). Only twelve out of twenty-three infants preferred the disfluent passages. This finding suggests that despite their greater disfluency overall, ten month olds are not sensitive to the prosodic cues to disfluency in adult-directed speech.

By contrast, our preliminary results with the fourteen 22 month olds suggests that they prefer the fluent speech stimuli (8.3 s) to the disfluent stimuli (6.9 s). However, with this small number of participants to date, the difference is not significant.

5.6. Follow-up studies

We are currently gathering more data to determine whether this difference in the older infants with the adult-directed stimuli is in fact significant, indicating that infants at this age differentiate disfluent from fluent utterances. If so, a second experiment using low-pass filtered speech will help to determine whether the older infants are using prosodic as well as grammatical cues to disfluency. Addtionally, it will interesting to run a comparable study with this older age range using infant-directed stimuli, to determine whether there is in fact a difference between the younger infants and older infants in how well they detect disfluency in adult- versus infant-directed speech.

6. Discussion

The current study examined two contexts for disfluency in maternal speech: One, a large, infant-directed corpus, in which speech samples were collected in the home environment in relative quiet, and the other a smaller sample collected in the chaotic environment of an airport waiting room, in which the mother was interacting with strangers as well as her own infant. This latter sample contained a much higher number of disfluencies than the more infant-directed corpus. However, ungrammatical word sequences were reliably marked by prosodic disfluency.

We next examined the extent to which these prosodic cues are detectable by adult raters and infants in a behavioral test. Disfluent and fluent utterances were highly discriminable by adult raters, especially the adult-directed speech stimuli. However, the picture with the infants is more complex.

In an initial experiment with infant-directed speech samples, one out of two groups of infants showed a significant difference in their preferences for the fluent and disfluent samples, but the other group did not. In the group showing a preference, the preference was correlated with the disfluency scores given to the samples by adult raters. Overall, these results are suggestive that infants are able to differentiate fluent and disfluent speech in infant-directed speech, but clearly further research is needed.

We next examined whether infants are able to differentiate fluent and disfluent utterances in adult-directed speech. Despite the more salient disfluencies in these samples - at least, according to adult ears - the ten month olds did not show any differences in their listening preferences. By contrast, an ongoing study with an older age group, 22 month olds, found a preference for the fluent passages.

Overall, these findings suggest that infants are indeed sensitive to the presence of disfluencies in speech, but the characteristics to which they attend may vary across different ages. Twenty-two month olds are well on their way to grammatical competence, and may even be combining words productively. Therefore, these older infants may detect disfluency based on properties of the language at which they are already competent, such as the repetition of function words, as well as the overall prosodic properties of the utterances. These two types of cues together may allow older infants to exclude from grammatical analysis both utterances and prosodic breaks which are ill-formed due to disfluency. On the other hand, younger infants are entirely reliant on the prosodic characteristics of the speech signal, and their ability to determine whether an utterance or prosodic break is a legitimate source of information about the grammar is limited to this aspect of the speech stream.

If so, these younger infants might be limited to infant-directed speech as a source of *reliable* input, while older infants may be able to make use of prosodic and structural information in both infant- and adult-directed speech -- infant-directed speech because the prosodic information is highly reliable, and adult-directed speech because the infants can detect disfluency, and reject it as input. This suggests a trade-off. On the one hand, infant-directed speech contains less prosodic information about grammatical structure, because it is shorter and less complex. On the other hand, adult-directed speech contains a greater wealth of information (and constitutes the majority of the ambient linguistic input), but these data are less reliable as input. This might suggest that infants begin by paying attention to the simpler, infant-directed speech, and only later (though well before they attain productive competence) access the adult-directed ambient input.

An additional puzzle lurking in the background of this discussion is: Why is infant-directed speech fluent and error-free? On first glance, the answer is simple. Mothers wish to be understood by their infants and therefore speak more precisely and fluently. However, this idea doesn't really hold up. Normal speakers don't "choose" to make speech errors and disfluencies in every day conversation. Speech errors are just that - errors. Most of the time, we are even unaware that we are producing them. Therefore, how can mothers choose to *not* make speech errors? One possibility is that because infant-directed speech is both slower and shorter than typical adult-directed speech, there are simpler fewer chances for errors. However, there is another reason for disfluency in adult conversational speech - the interaction of the second speaker. Conversational speech is a process of give and take - a speaker monitors in an on-going fashion the behavior of the listener to determine whether they are being understood. Spoken and gestural activity on the part of the listener can actively alter the intended output of the speaker - thereby creating disfluency. Many of the disfluencies generated in the transcript from MOT4 were of this type. MOT4 was clearly responding to head nods and gestures of comprehension from the listener, and interrupting herself to respond to overlapping speech by her conversational partners. In infant-directed speech, the partner is less active. Mothers may respond to the behavior of their infants, but since infants by definition do not understand their mother's speech, they are unlikely to respond in a way that will alter the mother's intended speech output.

Our answer to Chomsky's [1] and Newport et al.'s [8] contradictory assertions is that they are both right. The speech input environment of the child (and the infant) is complex and varied, both within a child's, and across different children's, experience. This input contains some instances of simple, error-free highly fluent utterances targeted directly to the infant in a quiet home environment. And it contains many noisy, highly disfluent, complex utterances for which the infant is simply along for the ride. We have only just begun to ask to what extent each of these types of input contributes to the language acquisition process. The current study suggests that infants may at least bring some tools to the table to help them differentiate "good", fluent input, from "bad", disfluent input.

7. Acknowledgements

The authors are extremely grateful for the long-term participation of the three mothers in the transcription study. We wish to thank the many participants in the preference experiment, as well as our adult raters. Thanks also to three undergraduate RAs, Rina Foygel, Megan Blossom and Greg Fay, for their work in transcribing and preparing speech samples, and to Katherine White for comments on this paper.

8. References

[1] Chomsky, N. 1962. Explanatory Models in Linguistics. In *Logic, Methodology and Philosophy of Science*, Ernest Nagel, Patrick Suppes, and Alfred Tarski (Eds.), pp. 528-550, Stanford, CA: Stanford University Press. (cited from Thomas, 2004)

[2] Fernald, A. 1985. Four-month-olds prefer to listen to motherese. *Infant Behavior and Development, 8,* 181-195.

[3] 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.

[4] MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

[5] Mandel, D.R., Jusczyk, P.W. & Kemler Nelson, D.G. 1994. Does sentential prosody help infants organize and remember speech information? *Cognition, 53,* 155-180.

[6] 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.

[7] 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.

[8] 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: language input and acquisition*, pp. 109-149. New York: C.U.P.

[9] 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.

[10] Soderstrom, M., Kemler Nelson, D.G., & Jusczyk, P.W. 2005. Six-month-olds recognize clauses embedded in different passages. *Infant Behavior and Development, 28,* 87-94.

[11] Soderstrom, M., Blossom, M., Morgan, J.L., & Foygel, I. in preparation. *The speech environment of 6-9 month old infants*

[12] Weijer, J. van de (2002). How much does an infant hear in a day? In J. Costa and M. João Freitas (Eds.). *Proceedings of the GALA2001 Conference on Language Acquisition*, pp. 279-282. Lisboa: Associação Portuguesa de Linguistíca.