J. Child Lang. **35** (2008), 869–902. © 2008 Cambridge University Press
doi:10.1017/S0305000908008763 Printed in the United Kingdom

Acoustical cues and grammatical units in speech to two preverbal infants*

MELANIE SODERSTROM
Brown University

MEGAN BLOSSOM
University of Kansas

RINA FOYGEL AND JAMES L. MORGAN
Brown University

(*Received* 3 *August* 2005. *Revised* 17 *May* 2007)

ABSTRACT

The current study examines the syntactic and prosodic characteristics of the maternal speech to two infants between six and ten months. Consistent with previous work, we find infant-directed speech to be characterized by generally short utterances, isolated words and phrases, and large numbers of questions, but longer utterances are also found. Prosodic information provides cues to grammatical units not only at utterance boundaries, but also at utterance-internal clause boundaries. Subject–verb phrase boundaries in questions also show reliable prosodic cues, although those of declaratives do not. Prosodic information may thus play an important role in providing preverbal infants with information about the grammatically relevant word groupings. Furthermore, questions may play an important role in infants' discovery of verb phrases in English.

INTRODUCTION

Children have generally been credited with knowledge about the grammatical structure of their language in the second year of life, based on their developing ability to combine words meaningfully in their own

[*] The authors wish to thank the two mothers and their families for their time and commitment to the study. This work was supported by a Kirschstein NRSA postdoctoral research fellowship 5F32HD042927 to MS and an NIH grant 1RO1HD32005 to JLM. We thank the reviewers for insightful comments on previous drafts. Address for correspondence: Melanie Soderstrom, University of Manitoba, Department of Psychology, P404 Duff Roblin Building, Winnipeg, Manitoba R3T 2N2, Canada.

productions. Because of this focus on children's productions, research on maternal speech as a source of linguistic input to children has also generally focused on speech to children who are themselves producing language. By contrast, recent perceptual studies with infants suggest that the roots of grammatical knowledge begin much earlier than children display in their productive behavior. For example, children often do not show productive use of function words and verb inflections until well into their third year of life. However, studies examining infants' perceptual preferences for listening to grammatical sentences compared with sentences which have been altered by replacing or scrambling words have found that sixteen-month-olds are sensitive both to the proper locations of functions words (Shady, 1996) and to the presence of verbal inflections (Soderstrom, White, Conwell & Morgan, 2007). Infants as young as eleven months detect when function words are replaced with nonce words (Shady, 1996). These and other findings suggest that grammatical knowledge begins to develop much earlier than previously thought. It is therefore important to examine the role that the characteristics of speech to preverbal infants (i.e. infants in the first year of life, who are not yet showing productive linguistic capabilities) might play in the development of grammatical knowledge. Such a role may manifest directly, in terms of the grammatical structure and complexity of the input, and/or indirectly, through extralinguistic cues such as the acoustical/prosodic organization of the input.

Intuitively, one might expect there to be large differences in the grammatical structure of speech to infants before and after the onset of productive language, at about twelve months. After all, our conversational interactions with partners who can respond in a linguistic fashion should be qualitatively and quantitatively different from those with preverbal ones. However, quantitative differences, at least, are difficult to find at this point in development.

One of the most salient and well-noted aspects of maternal, infant-directed (ID) utterances is that they are relatively short and less complex compared with adult-directed (AD) speech. While AD speech has a mean length of utterance (MLU) of eight or more morphemes, speech directed to children across a variety of ages is consistently about half that. Some increases in complexity of maternal speech have been found at a later point in development, at or after twenty-four months (Nelson, 1973; Phillips, 1973; Kaye, 1980; Stern, Spieker, Barnett & MacKain, 1983), but comparisons before and after twelve months have found little evidence of quantitative changes (Phillips, 1973; Snow, 1977). It is possible that such differences in complexity may also take place at earlier ages, but even there the evidence is sparse. Papousek, Papousek & Haekel (1987) found greater simplicity in an analysis of speech to three-month-olds compared with speech to older infants in other studies, but did not examine older infants within their study. Similarly, Kaye (1980) found greater simplicity in speech to infants younger than six months compared with the same infants at two years, but did not examine intervening ages. One study (Sherrod, Friedman, Crawley, Drake & Devieux, 1977) actually found shorter utterances at eight months than four months.

While quantitative measures have found little differences between speech to preverbal infants and those beginning to produce speech themselves, some of the same researchers have noted qualitative differences, such as very long or whispered utterances in speech to very young infants (e.g. Phillips, 1973; Snow, 1977). For example, Phillips noted that when speaking to an eight-month-old, a mother 'appeared to be talking to and for herself, in spite of the childish intonation patterns and onomatopoetic sequences she often used' (p. 184). Snow, in a more detailed examination of the characteristics of speech to preverbal infants, suggested that the change toward a conversational interaction between mother and infant occurred slightly earlier, by about seven months. Other qualitative differences have also been found. Kaye (1980) found a large number of isolated one-word greetings or phatic utterances like *yeah* or *sure* in speech to preverbal infants. An analysis of the type of speech to infants found an increase from three to six months in the number of informative utterances compared with those whose purpose was to communicate affect (Penman, Cross, Milgrom-Friedman & Meares, 1983).

Such qualitative differences suggest that there may be subtle but important differences in the grammatical properties of speech to preverbal infants compared with young children beginning to access linguistic knowledge. Phillips (1973) found tentative evidence for greater variability in speech to eight-month-olds than to much older infants. To our knowledge, this idea has not been pursued further. However, it is supported by the above findings of longer, more adult-directed utterances (Phillips, 1973; Snow, 1977), together with much shorter, quasi-linguistic or phatic utterances. These differences suggest that in order to understand the grammatical nature of linguistic input to preverbal infants, it is important to examine not merely mean complexity of speech across a sample, but rather the relative amounts of utterances of varying complexity, from isolated phatics to complex, multi-clause utterances. One goal of the current study is to examine the relative amounts of maternal input from utterances of varying degrees of complexity.

Another important question regarding the syntactic structure of the input to young infants has to do with the frequency of questions. Speech to preverbal infants, like that to older infants, has generally been found to contain a large percentage of interrogatives, around thirty percent of utterances (e.g. Snow, 1977). However, one longitudinal case study of speech to a Dutch-learning preverbal infant found a much smaller number of questions, about twelve percent (van de Weijer, 2002*a*). Questions, particularly *yes*/*no* questions, are both acoustically and structurally different from declarative forms, and may play a unique role in the input. We will therefore examine in this study the extent to which questions play a role in the input to preverbal infants.

Finally, we will examine the role that the acoustic characteristics of speech to preverbal infants may play in grammatical development. Numerous researchers have theorized that that acoustic or prosodic characteristics of speech may help infants to organize the speech they are hearing into grammatically relevant units, often referred to as 'prosodic bootstrapping' (e.g. Morgan & Newport, 1981; Gleitman & Wanner, 1982; Jusczyk, 1998). However, this idea has drawn critics, in part because of the existence of prosody–syntax mismatches (Pinker, 1994; Fernald & McRoberts, 1996).

Compared with the sparse literature on the grammatical properties of speech to preverbal infants, a number of researchers have characterized the unique prosodic and acoustic properties of ID speech, including higher pitch, greater variability in pitch and longer pauses than typical AD speech (e.g. Stern *et al.*, 1983; Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies & Fukui, 1989). Fernald *et al.* (1989) found that these characteristics are manifest cross-linguistically and in both mothers' and fathers' speech, although there are perhaps differences in degree. ID speech is also slower than AD speech, and (at least in English) contains special intonational patterns not found in AD speech (Sachs, 1977).

Prosodic cues to syntactic boundaries are widely attested in AD speech (e.g. Klatt, 1975; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992). In the laboratory setting, research has found sensitivity to these cues by two months of age (Mandel, Jusczyk & Kemler Nelson, 1994) and evidence of their use in fluent speech processing by six months (e.g. Soderstrom, Seidl, Kemler Nelson & Jusczyk, 2003).

The findings that six-month-olds use clause- and phrase-level prosodic cues to parse speech in the laboratory setting suggest that these cues might play a crucial role for infants who are in the process of discovering the grammatical characteristics of their language. The speech that is used in these laboratory experiments, however, is necessarily a poor representation of the kind of speech that infants hear in the real world. For example, reliable differences have been found in the presence and distribution of prosodic boundaries in read versus spontaneous speech (Blauww, 1994). The relationship between prosody and syntax in the real world may be much messier than that in carefully controlled laboratory experiments.

There are reasons to think that infant-directed speech, in particular, may serve to reduce prosodic cues to syntax. First, as previously mentioned, ID utterances are on average much shorter than AD utterances. Therefore, while cues to utterance boundaries may be strong, it is not clear whether infants hear utterances long enough to be organized in multiple prosodic units. Second, ID speech tends to contain a large number of pronouns, rather than full noun phrase subjects. Sentences with pronominal, rather than full NP, subjects tend to be organized differently in terms of their prosodic structure (Gee & Grosjean, 1983), and this organization may not match the syntactic structure of the sentence as closely.

On the other hand, there are also reasons to propose that infant-directed speech may enhance prosodic cues to syntax. The prosodic characteristics of ID speech described above may be viewed as exaggerations of the AD form – therefore prosodic cues that exist may be more salient. Some behavioral studies have found evidence that young infants were more sensitive to prosodic cues to phrase boundaries in ID speech than AD speech (e.g. Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward & Piwoz, 1992).

While we know a great deal about the perceptual sensitivities of six-month-olds, we know much less about how prosody and distributional/syntactic properties interact in the speech they are hearing. There is at least suggestive evidence in one study that one prosodic cue to clause boundaries (pre-boundary vowel lengthening) may be stronger in speech to preverbal infants (between nine and thirteen months) than to older infants with developing productive vocabularies (Bernstein Ratner, 1986). A more recent study analyzed speech to slightly older infants (at approximately 1;2). Fisher & Tokura (1996*a*), examining maternal speech to a small number of American English and Japanese learning infants, found strong prosodic cues to utterance/clause boundaries in both languages (pause duration, pre-boundary vowel lengthening and pitch change). Fisher & Tokura also examined the phrase-level boundary between subject and verb phrase, but results were less consistent. Only the most sensitive measure found any statistically significant differences in the various acoustical measures between phrase-boundary and non-boundary locations, and these acoustical measures were different between the two languages. For English, there was a significant effect of syllable duration, while for Japanese, the significant effect was in pitch lowering. Fisher & Tokura (1996*b*) propose that infants may use prosody to access phrases only indirectly, by detecting utterances which are phrase-level sentence fragments.

There are a number of issues left unresolved by this study. First, as noted by Fernald & McRoberts (1996), this study and those of speech to older infants and young children do not distinguish carefully between 'clause' and 'utterance'. This is important because while utterance boundaries are clearly very well marked prosodically, it is possible that clause boundaries PER SE are not. Second, the extent to which phrase boundaries within a clause are marked needs more careful examination. The measure with which Fisher & Tokura found an effect required multiple syllables preceding the boundary. This may occur either because the subject itself is multisyllabic, or in the case of questions because an auxiliary is placed before the subject. Since Fisher & Tokura did not distinguish between declarative and question utterances, it is not clear whether this effect was being carried by questions, or by multisyllabic declarative subjects. Finally, given the behavioral data suggesting that very young infants are sensitive to prosodic characteristics of speech, and the finding in the Bernstein Ratner study (1986) suggesting that these cues might be stronger in speech to YOUNGER infants, it is important to consider whether these cues exist in speech to infants younger than twelve months, who are truly preverbal.

In the current study, we examine the characteristics of the speech input to two infants between six and ten months, with follow-up visits at twelve and eighteen months. First, we will characterize the grammatical properties of the utterances themselves. Then we will analyze the prosodic characteristics that may provide information to the infants about the grammatical organization of their language. We are interested in the following questions: What are the relative amounts of utterances of varying degrees of complexity, such as phrasal fragments or longer utterances containing multiple prosodic units? What role might questions play in the input? Do prosodic cues within the utterance exist in speech to preverbal infants?

METHOD

Participants

The participants in this study were two mothers (MOT1 and MOT2) with young male infants, INF1 and INF2 respectively. Both mothers were living in Rhode Island at the time of the study, but had lived various other places before settling there. Both of INF1's parents had earned graduate degrees and worked outside the home. While the parents were at work during the day, INF1 was in a daycare with other infants and children of various ages. INF1 had two other siblings, a three-year-old sister and a five-year-old brother. INF2's father had a graduate degree and worked outside the home, while his mother stayed home with INF2. INF2 had one brother, a five-year-old who was in school during the day.

Data collection

Recordings for MOT1 were collected via Azden wireless lapel microphones through a receiver onto a Sony digital IC recorder. Recordings for MOT2 were recorded through the Azden lapel microphone plugged directly into the Sony recorder. These files were converted to wav files and then transcribed using the CLAN system (MacWhinney, 2000). The equipment was left with the participants, who were asked to record an hour per week of regular

TABLE 1A. *File length and starting ages for MOT1's transcripts*

File	Total length	Starting age (months.days)
1	0:30:52	5·29
2	0:10:16	6·12
3	0:45:48	7·03
4	0:46:31	7·17
5	1:08:17	8·01
6	0:10:38	8·07
7	1:09:39	8·17
8	0:44:40	8·24
9	0:53:00	9·01
10	1:21:28	9·15
11	0:54:38	9·25
12-1	1:01:13	12·21
12-2	1:04:02	12·29
18-1	1:15:46	18·07
18-2	1:10:14	18·10

interaction with their infants between the ages of six and ten months. Follow-up sessions were recorded at twelve and eighteen months for comparison. Additionally, at eighteen months, video recordings were made of the sessions, and microphones were put on both the mother and infant. Both mothers were asked to record during a time when they would normally be interacting with their infant, but to minimize the presence of other adults and siblings. MOT2 was more successful at minimizing the presence of additional participants than MOT1, since she was alone with the infant during the day. In MOT1's recordings, the siblings were present at least some of the time during each of the recording sessions. This difference in the recording environments may have had consequences for the character of the utterances produced by the two mothers, and will be discussed later.

Tables 1a and 1b list the files and ages at which they were obtained. From MOT1 we obtained twelve recording sessions between six and ten months, including two short ten-minute sessions (files 2 and 6) due to equipment failure. The remaining sessions averaged forty-nine minutes in length. We also combined a half-hour session by MOT1 with a longer session the following day (file 10) to make a final set of eleven files. From MOT2, we obtained a number of recordings of a variety of lengths during this time period, due in part to the stopping and starting of recording equipment during a session. Recordings from a given week were combined into a single file, and recordings shorter than five minutes that were not immediately following or followed by a longer recording were excluded from the analysis. This generated fifteen files for analysis for MOT2 between six and ten months. Two additional sessions were collected for each infant at twelve and eighteen months.

TABLE 1B. *File length and starting ages for MOT2's transcripts*

File	Total length	Starting age (months;days)
1	1:04:09	6·15
2	1:07:09	6·23
3	0:59:42	6·29
4	0:32:13	7·08
5	0:27:23	7·12
6	0:52:47	7·19
7	1:24:14	7·26
8	0:37:18	8·05
9	0:59:19	8·09
10	0:29:47	8·18
11	1:23:44	8·23
12	0:40:38	9·0
13	0:45:03	9·09
14	1:19:13	9·15
15	1:03:33	10·11
12-1	0:31:44	12·17
12-2	1:05:53	12·28
18-1	1:05:58	17·22
18-2	1:02:25	18·01

Files were transcribed using the basic CHAT protocol, and coded for a variety of syntactic and prosodic properties (described below) on a separate coding tier. Each transcript was transcribed and coded initially by a primary transcriber. The entire sound file and transcript were then reviewed by a checker. Differences of opinion between the transcriber and checker were resolved by reviewing the material together. A total of 3,675 utterances (13% of all utterances, which included two files from each mother) were then re-coded from scratch in order to obtain measures of reliability. Reliability was above 90% for all code types.

MacArthur-Bates Inventories

During the first eighteen-month recording session, the participant was asked to fill out a MacArthur-Bates Communicative Development Inventory: Words and Sentences (Fenson *et al.*, 1993) for the child and a short questionnaire about the infant's home life. INF1 and INF2 had very similar verbal production at eighteen months, scoring just under the 25th percentile in words produced. Additionally, neither infant combined words or used inflectional morphemes by eighteen months. The similarity in their verbal development at eighteen months was unplanned, but interesting given that the speech environments of the two infants seemed to be very different.

Utterances

An initial difficulty in examining the relationship between prosody and syntax in speech concerns determining what constitutes an utterance. Spontaneous conversation includes sentence fragments and ellipses, and prosodic considerations factor into our perceptions of utterance boundaries. Previous work has variously used strict acoustical criteria such as pauses greater than 300 ms (e.g. Stern *et al.*, 1983), grammatical units or transcription markers such as commas and periods (Fisher & Tokura, 1996*a*) as indicators of utterance boundaries. Because fluent speech is extremely variable and described by a variety of prosodic factors, strict acoustical criteria, while being objective, are likely not as sensitive or accurate as more subjective methods. Furthermore, these strict acoustical criteria, while more straightforward, may seriously underestimate the length and complexity of ID utterances, especially considering the exaggerated prosodic boundaries produced in ID speech.

In our transcripts, we used the best judgment of the transcribers to determine utterance boundaries. In cases where utterances were not separated by long silent pauses, utterances boundaries were determined based primarily on prosodic, rather than syntactic or semantic considerations, because from the perspective of the preverbal infant, these are most salient. In a few cases where semantic/discourse considerations suggested an utterance boundary, but the words clearly formed a single prosodic unit, the speech bout was treated as one utterance. In other cases, a very strong prosodic break and/or long silent pause overrode the clear semantic and syntactic connection between two speech bouts, and they were treated as separate utterances. These cases were rare. A more common difficulty was deciding when two prosodically and syntactically related yet separate clauses or phrases should be treated as one or two utterances, e.g. *Oh, that's nice*. (one utterance with a prosodic break) vs. *Oh*. *That's nice*. (two smaller utterances, neither of which contained an internal prosodic break). Such considerations were decided upon on a case-by-case basis, by agreement of the transcriber and checker, based on whether the speech sounded like it was intended as one utterance or two. To the extent that inconsistencies existed, this reflected the fact that such boundaries fall in a continuum, rather than error on the part of the transcribers. That said, there was relatively good agreement on the assignment of utterance boundaries during the checking process.

Coding and acoustical analyses

We first examined the overall syntactic complexity of the speech to these two infants, by assigning each utterance a category code. There were two categories of utterance excluded from additional analyses. Unclear or untranscribable utterances were excluded from all further analysis, while routines (jingles, songs, book reading, etc.) were excluded from analysis except where noted. However, PARALINGUISTIC vocalizations (vocalizations with phonological content but lacking a clear lexical meaning) were included. Non-linguistic mouth noises such as involuntary laughter, sneezing, coughing, etc., were not counted in the analysis. However, some 'laughter' was included, if it seemed voluntary and communicative.[1] The remaining utterances were classified according to their syntactic/linguistic complexity. Isolated vocal sounds with no semantic content (generally vocal play such as raspberries and babbling) were classified as NONSENSE utterances. Isolated linguistic sounds with questionable semantic status were classified as QUASI-WORD utterances. Any utterance that contained at least one fully lexical word, but did not contain any phrase-level constituent structure, was classified as a WORD utterance. Since the percentage of words that infants hear in complete isolation is of interest in understanding how infants find individual words in fluent speech, we also counted the subset of these word utterances which were a single full word in isolation. The remaining three categories examined the complexity of multiword utterances. A PHRASE utterance contained some constituent structure, but did not contain both a subject and predicate. Note that this utterance category included elided sentences and imperatives. An imperative, while a grammatically complete clause, does not provide the infant with full constituent information at the level of the sentence, or IP. This definition provided a conservative estimate of the syntactic complexity of the input to the infants compared with other studies. We reserved the designation of CLAUSE utterance for any utterance which contained a full subject–predicate combination. Finally, an utterance was considered a MULTICLAUSE utterance only if it had at least two predicates and at least one subject. Appendix 1 provides a table outlining these categories with examples. Reliability for this measure was above 95%.

We then examined the major prosodic breaks, i.e. significant pausing and/or pitch changes, within utterances, that had not been judged to be utterance boundaries (generally, these were marked by a comma, a hash mark to indicate a pause or a marker of disfluency according to the CHAT protocol). As with judgments regarding utterance boundaries, the transcription of these prosodic breaks was necessarily subjective. Given the more subtle prosodic cues associated with phrase boundaries than utterance boundaries, it was expected that reliability would be lower for this measure. While

[1] These judgments were all necessarily subjective. Care was taken to reach agreement between the original coder and the checker, and such agreement was high. To the extent that these judgments may have been unreliable, the effect would have been a small over- or underestimation of the proportion of utterances in the NONSENSE and QUASI-WORD categories.

transcriber reliability for these boundaries was still relatively high for MOT2 (87%), it was much lower for MOT1 (73%), whose input was overall more adult-like. Again, the lower reliability in these judgments reflects the fact that these cues fall along a spectrum – different coders were more conservative or liberal in judging whether a set of prosodic cues constituted a major break.[2] These breaks were categorized as fluent, disfluent or other (see Appendix 1). These intra-utterance units separated by a prosodic break were then coded for syntactic complexity as described for utterance complexity. For example, the phrase-level utterance *Sit up, please, the right way*. was coded as containing two phrases and a prosodically isolated word. The utterance *Yes, please*., a word-level utterance, was coded as containing two prosodically isolated words.

To examine quantitatively whether acoustic properties of the speech stream reliably signal syntactic boundaries, one transcript from MOT1 (file 11, when INF1 was 0;9.25) and two transcripts from MOT2 (files 13 and 14, when INF2 was 0;9.9, and 0;9.15 respectively) were selected for acoustical analysis. We selected transcripts from when the infants were nine months old because the mothers were by this time very comfortable with the recording sessions, but the infants were still well within the 'preverbal' period. Many of the behavioral studies showing sensitivity to prosodic cues to syntax have focused on nine-month-olds. Acoustic analyses were performed using Praat software (Boersma, 2001) by a trained analyst.

We first took approximately 250 maternal syllables (248 for MOT2, 251 for MOT1) near the beginning of two transcripts, MOT1 file 11, and MOT2 file 14. Nonsense and quasi-words were excluded from this analysis. To increase the numbers in the categories of interest (clause- and phrase-level boundaries), we then targeted specific utterances that contained phrase- and clause-level boundaries to include in the sample from the remainder of the original two files, and also included additional utterances from file 13 for INF2. Our final sample contained 803 syllables from INF1 and 666 syllables from INF2. We measured the following four acoustical properties:

Pause: We measured the silent pause (if any) after each syllable until either mother or baby vocalized again. Pauses longer than two seconds (usually at an utterance boundary) were excluded from analysis.

[2] Although coders were explicitly instructed to attend only to the prosodic information in making such judgments, grammatical knowledge on their part may have contributed to overestimates of the number of such prosodic boundaries at grammatical phrase and clause boundaries. However, because our acoustic measures, which were done across all syllables (irrespective of the transcriber coding of boundaries), found significant prosodic cues to utterance-internal boundaries, we can be confident that these cues do exist.

Vowel duration: We measured the length of each vowel, basing our judgment of the onset and offset of the vowel on an analysis of the waveform and spectrogram, as well as audio playback. Where an exact boundary between vowel and consonant could not be determined (for example, in some cases where the adjacent consonant was a liquid), the midpoint of the ambiguous region was chosen. We then computed difference scores between the durations of the pre- and post-boundary vowels.

Change in pitch: We found the peak pitch value of each vowel and computed difference scores between the pre- and post-boundary vowels. Vowels with undefined pitch according to the Praat software were excluded from analysis. This difference was converted to semitones.

Intensity: We also recorded the average intensity across the vowel, and computed difference scores between the pre- and post-boundary vowels.

Additionally, we coded each utterance as to whether it was (or contained) a question. Questions were subdivided into *yes/no* and *wh*-forms. We further categorized these questions according to whether they contained cues to their status as a question based on their prosodic characteristic or inversion, or both. Since transcriber coding of prosodic cues to questions might be unreliable, listener ratings and acoustic analyses were also performed on file 11 of MOT1 by a trained analyst to determine whether prosodic cues were detectable (described in more detail in the results section).

RESULTS

Our findings are presented in three sections. In the first section, we describe some general characteristics of the speech samples from the two mothers. In the second section, we describe the transcriber coded qualitative properties of the utterances. These data address primarily our first two questions: What are the relative amounts of utterances of varying degrees of complexity, such as phrasal fragments or longer utterances containing multiple prosodic units? What role might questions play in the input? This section is divided by grammatical category type, with separate descriptions for clauses, phrases, words and paralinguistic utterances. For each sub-section, we also describe the extent to which our transcriber-coded prosodic boundaries isolated that type WITHIN utterances. Our analysis of questions within our samples is also provided in a separate subsection. In the final section, we provide a quantitative analysis of prosodic characteristics of phrase-, clause- and utterance boundaries, in order to address the third question in more detail: Do prosodic cues within the utterance exist in speech to preverbal infants?

TABLE 2. *General characteristics of speech samples: the number of utterances, complexity of speech (measured by MLUw) and speech density (measured by words and by utterances)*

MOT/Age	Number of utterances Total[a]	Number of utterances ID	MLUw ID	Word density total (words/min)	Word density ID (words/min)	Utterance density total (utterances/min)
MOT1						
6–10 months	9067	6926	4·376	83·97	77·04	17·58
12 months	2297	1920	4·580	89·54	84·26	18·41
18 months	2266	2224	4·141	65·22	64·35	15·57
MOT2						
6–10 months	10 904		2·949	42·00		13·97
12 months	1466		3·440	52·48		15·03
18 months	1473		3·507	41·50		11·63

[a] Number of utterances, and word density in words per minute are reported for all of the mother's utterances (total) and the mother's infant-directed utterances only for MOT1 (ID). MLU in words is reported only for infant-directed utterances (both mothers), while the utterance density in utterances per minute is reported for all of the mother's utterances.

General characteristics of speech samples

Table 2 describes the number of utterances, MLUw (mean length of utterance in words) and speech densities in each transcript. For the density of speech, we examined both the number of utterances per minute, and the number of words per minute. Overall, MOT1 produced longer utterances, and more input per unit time, in both utterances and words, than MOT2 (there was a greater total number of utterances for MOT2 because she had a longer amount of total recording time). Some of these utterances of MOT1 were directed at older siblings and other adults; however, even if we look only at her speech to INF1, the difference is striking. In fact, removing the AD and CD utterances only lowers the word density by 7 words per minute (from Table 2: 83·97–77·04). We surmise that the mere presence of other adults and older children may have had an effect on MOT1's speech, serving to make it more adult-like, just as one might tend to produce characteristics of ID speech in the presence of an infant, even if one is addressing another adult. By contrast, MOT2 was alone with her infant for most of the day, and her recordings were produced almost exclusively in isolation with her infant. Although the pattern of results presented below changed very little whether we examined all maternal utterances or only those that were infant-directed, we excluded all non-ID utterances from analysis for MOT1 in subsequent tables to make them more comparable with those of MOT2.

TABLE 3. *Proportions of utterance types for MOT1 and MOT2*[a]

MOT/Age	Non-sense	Quasi-word	Word: [b]*Isolated*	Word: *All*	Phrase	Clause	Multi-clause	Unintel-ligible	Routine
MOT1									
6–10 months	0·11	0·12	0·07	0·10	0·13	0·38	0·15	0·02	0·00
12 months	0·04	0·09	0·09	0·12	0·14	0·44	0·18	0·01	0·00
18 months	0·01	0·11	0·10	0·12	0·15	0·45	0·16	0·01	0·00
MOT2									
6–10 months	0·12	0·18	0·15	0·18	0·15	0·28	0·06	0·02	0·01
12 months	0·04	0·20	0·10	0·14	0·16	0·30	0·07	0·01	0·08
18 months	0·03	0·22	0·15	0·19	0·15	0·28	0·09	0·00	0·04

[a] See Appendix 1 for descriptions of the categories.
[b] The category Word: *All* reports utterances with word-level complexity (i.e. does not contain any multiword phrases). The category Word: *Isolated* reports only those utterances containing a single word in isolation. The *Isolated* category is a subset of the *All* category. Apart from this, utterance types are mutually exclusive, and a given utterance was classified into one and only one of these categories.

Grammatical units

Our primary goal was to determine what information is available to infants about the grammatical structure of their language. Table 3 describes the proportions of utterances of each category produced by our two mothers. We did not see significant developmental changes within the six- to ten-month time period, so the data presented were collapsed across this age range. Changes at the twelve- and eighteen-month sessions were also relatively few, and will be noted as appropriate below. Overall, both mothers produced a variety of short and long utterances. MOT1 produced a larger number of clauses than MOT2, while MOT2 produced more single-word utterances. These data are discussed in more detail in the following sections.

To examine the extent to which prosodic information might help the infant to find grammatical structures, we first examined a qualitative measure, the utterance-internal 'major prosodic breaks' coded by the transcribers. As expected, we found that fluent prosodic breaks far out-numbered disfluent and other prosodic breaks. While on average there was only 1 prosodically marked disfluency in 100 utterances, fluent prosodic breaks occurred in MOT1's speech at a rate of 32 in 100 utterances, and in MOT2's speech at a rate of 22 in 100 utterances. Major prosodic breaks are generally fluent in the input, and might therefore provide a reliable source of information to infants about units of speech. Table 4 presents an analysis of the additional grammatical units available to the infants if these utterance-internal major prosodic breaks are taken into account. These additional grammatical units were coded using the same criteria as

TABLE 4. *Additional grammatical units generated by considering utterance-internal major prosodic breaks*

	Nonsense/ Quasi-word	Word	Phrase	Clause	Total
MOT1[a]	603 (27%)[b]	838 (37%)	373 (17%)	444 (20%)	2258 (100%)
MOT2	1009 (42%)	888 (37%)	235 (10%)	256 (11%)	2388 (100%)

[a] Reliability for coding the presence of a fluent break for MOT1 was low (73%). Reliability for coding the categories of the transcribed units for both mothers was above 90%.
[b] Data are reported by total number and in parentheses by the percentage of the total additional grammatical units. Only 6–10-month-old data are shown.

TABLE 5. *Clause types*

MOT/Age	1[a]	2	3	4+	NF[b]	F	Emb[c]	Conj
MOT1								
6–10 months	0·71	0·22	0·05	0·01	0·16	0·84	0·49	0·51
12 months	0·71	0·21	0·06	0·02	0·18	0·82	0·60	0·40
18 months	0·74	0·21	0·04	0·01	0·19	0·81	0·62	0·38
MOT2								
6–10 months	0·83	0·14	0·03	0·00	0·17	0·83	0·62	0·38
12 months	0·82	0·15	0·03	0·00	0·15	0·85	0·81	0·19
18 months	0·76	0·18	0·05	0·01	0·15	0·85	0·71	0·29

[a] The first section reports the percentages of full clause utterances (i.e. containing at least a subject and predicate) by the number of clauses.
[b] The second section reports the percentages of non-finite (NF) and finite (F) non-matrix clauses in these utterances.
[c] The last section reports the percentages of embedded (Emb) and conjoined (Conj) non-matrix clauses.

for the utterances themselves, as described in the methods section and Appendix 1.

Clauses. As shown in Table 3, MOT1 produced 53% of her utterances as full clauses – 38% single clause and 15% multiclause at six to ten months. The amount of full clauses increased somewhat at twelve and eighteen months to about 62%. By contrast, MOT2 produced about 34% of her utterances as full clauses at six to ten months, 28% single clauses, and only 6% multiclause, and showed little or no increase at twelve and eighteen months. An examination of Table 4 shows that MOT1 also had a larger percentage of prosodically isolated utterance-internal clauses than MOT2 (20% vs. 11%). Table 5 presents a further analysis of the clause types in the speech by MOT1 and MOT2. Approximately three-quarters of the clause-level utterances are single clauses. Furthermore, in keeping with previous findings (Morgan, 1986), the majority of multiclause utterances contained only two clauses. A small percentage contained three clauses and very few contained four or more. Multiclause utterances contained somewhat higher numbers of embedded than conjoined non-matrix clauses across all three age ranges for MOT2, and at twelve and eighteen months for MOT1. Over 80% of these non-matrix clauses were finite.

Overall these results suggest that even young infants may be hearing a large percentage of complete sentences, and some rather complex utterances, particularly in an environment where siblings are present. Whether these sentences contain enough complexity that prosodic information becomes necessary and/or useful to the infant will be examined more carefully below.

Phrases. Fisher & Tokura (1996*b*) suggest that the presence of isolated phrases in maternal speech might provide infants with important clues to syntactic units. The two mothers in the current study had similar proportions of isolated phrase-level utterances – 13% for MOT1 and 15% for MOT2 (again, see Table 3). MOT1 had a slightly higher percentage of prosodically isolated phrases, 17% vs. 10% (see Table 4). Together with the data on full clauses, these data suggest that utterance-internal high-level prosodic information might indeed play a role in infants' organization of words into syntactically relevant units, as suggested by Fisher & Tokura (1996*a*).

Lexical words. Familiarity with isolated words might play an important role in the developing ability to find word boundaries (Brent & Siskind, 2001). Table 3 provides both the total number of utterances classified in the category 'word' as described in Appendix 1 (Word: ALL), and also the percentage of utterances that are completely isolated lexical words (Word: ISOLATED). The percentage of isolated single-word utterances for MOT1 was 7%; for MOT2 it was 15%. These results are similar to those of Brent & Siskind, who examined speech to English-learning infants between nine and fifteen months, and found approximately 9% isolated words.

Single words were also the most common unit isolated by the fluent prosodic breaks (37% for both mothers). This high-level prosodic information might also allow infants to detect individual words which might otherwise remain hidden in the speech stream.

Paralinguistic utterances and quasi-words. The characteristics of ID speech necessarily reflect the contexts in which they occur and the purposes that an adult might have in speaking to a preverbal infant. Our recordings were filled with gasps and sighs, little laughs and a variety of sound effects, which were clearly intended as purposeful communication on the part of the mother. While it might be tempting to exclude entirely these paralinguistic exclamations from the analysis, they often open or close an utterance, and on rare occasions are even inserted partway through an utterance. They are therefore an important part of the infant's exposure to linguistic information.

Studies have generally found the proportion of babble, fillers, social expressions or other paralinguistic utterances in infant-directed speech to be high. Van de Weijer (2002*a*) placed the number at about 49% in his study of a Dutch-learning infant, while Snow placed the number somewhat lower, between 20 and 40%. This difference likely reflects a difference in how utterances were categorized in the two different studies, as well as the contexts in which the data were collected. Unfortunately, there is no standard definition or set of utterance types in studies of either ID or CD speech. While this allows for flexibility in examining different characteristics of maternal speech, it makes it difficult to compare results across studies.

Our estimates of paralinguistic speech, collapsing across the categories of nonsense words, quasi-words and routines, were slightly lower than those of van de Weijer but approximately on a par with those of Snow: 23% for MOT1 and 31% for MOT2 at six to ten months. The average rate for MOT1 at twelve and eighteen months is even lower (13% and 12% respectively). Interestingly, this value for MOT1 is much lower even than the estimates found by van de Weijer for adult–child and adult–adult speech (31% and 39% respectively). Again, the extent to which such differences are due to individual differences, the specific activities and contexts of the MOT1 recordings or differences in how the categories were defined, is unclear.[3]

Clearly, paralinguistic speech is a major component of the input to preverbal infants. Such expressions might play a crucial role as carriers of affective information, but from the perspective of grammatical development, they might constitute a source of noise in the system. In particular, we may ask to what extent exclamations like *oh* and *hm* are acoustically distinguishable from lexical items, especially function words like *a* or *the*. When we examine the units of speech isolated by major prosodic breaks according to our coders, we find that a great deal of these breaks isolate nonsense words or quasi-words from the rest of the utterances, especially in MOT2's speech (27% for MOT1, 42% for MOT2). This might provide a valuable indication to infants that these words are not function words, which are unlikely to occur as prosodically isolated units. For young infants, prosody might also play a role in separating nonsense and semantically weak quasi-words from the rest of the utterance.[4]

[3] Of these paralinguistic utterances, approximately 10% for MOT1 and 20% for MOT2 were 'quasi-words'. Since this was a catch-all category, we further categorized the quasi-words from INF1 file1. Approximately 60% of these utterances were common interjections (*oh*, *hm*, etc.), another 30% were voluntary emotive interjections (gasps, laughs, etc.) and about 8% were sound effects.

[4] Again, coders' grammatical knowledge may have caused overall overestimates of the extent to which such prosodic breaks exist between paralinguistic speech and lexical words and phrases. We did not examine these paralinguistic exclamations acoustically in the current study.

TABLE 6. *Questions: proportions of* yes-no *and* wh-*questions marked by prosodic cues* (*pros*), *inversion* (*inv*) *or both*

MOT/Age	*yn*-pros	*yn*-inv	*yn*-both	*wh*-pros	*wh*-inv	*wh*-both	*wh*-only[a]	Total[b]
MOT1								
6–10 months	0·09	0·01	0·16	0·01	0·01	0·08	0·01	0·36
12 months	0·15	0·00	0·20	0·01	0·01	0·12	0·00	0·49
18 months	0·13	0·00	0·23	0·02	0·01	0·17	0·00	0·56
MOT2								
6–10 months	0·10	0·00	0·07	0·01	0·00	0·07	0·00	0·25
12 months	0·14	0·00	0·07	0·02	0·00	0·10	0·01	0·34
18 months	0·16	0·00	0·07	0·05	0·00	0·04	0·01	0·33

[a] Utterances marked only by the presence of a wh-word.
[b] Total proportion of question utterances (including utterances with tag questions) across all utterances in the corpus.

Questions. Questions are of particular interest in studying maternal speech because CD speech contains a large percentage compared with declaratives (Newport, Gleitman & Gleitman, 1977), and questions do not follow the simple SVO pattern of English. There has been considerable debate about the extent to which maternal speech patterns, like the percentage of *yes/no* questions, affect the acquisition of specific grammatical forms like auxiliaries (e.g. Newport *et al.*, 1977; Furrow, Nelson & Benedict, 1979). While one might not expect that infants of six to ten months would yet be learning about auxiliary use, Snow (1977) found even higher rates of questions in speech directed to very young infants (above 40% in three-month-olds). By contrast, van de Weijer found much lower rates of interrogatives overall, around 10% for ID speech, and 12% for AD speech in the Dutch infant's presence. In selecting stimuli for an analysis of the prosodic characteristics of questions, van de Weijer (2002*b*) found a greater number of *yes/no* questions than *wh*-questions in ID speech, but the reverse pattern in AD speech.

Our analysis (see Table 6) was consistent with the large number of questions found in the studies of English-learning infants, and a greater number of *yes/no* questions in ID speech – 26% (the sum of the proportions of the *yn*-pros, *yn*-both and *yn*-inv categories) of MOT1's utterances at six to ten months were or contained *yes/no* questions, compared with only 11% *wh*-questions (the sum of all four *wh* categories). The rates for both types at twelve and eighteen months were somewhat higher, up to 56% in the eighteen month samples, but preserved the higher ratio of *yes/no* questions. MOT2 had a somewhat lower rate of questions overall, but the pattern was similar – 17% *yes/no* questions at six to ten months compared with 8% *wh*-questions (again somewhat higher at twelve and eighteen months).

Counter to van de Weijer's finding with a Dutch-learning infant, a separate analysis of MOT1's AD-only speech found *yes/no* questions still in the majority at 60–70% of questions.

Our analysis also found very few examples of inversion without prosodic marking, while there were a great many instances of questions marked with prosody that did not contain overt inversion, despite our liberal criterion for inversion. An utterance was considered to be overtly inverted if it could not be produced fluently as a declarative, even if the auxiliary was missing. For example *He gonna play now?* would be coded as inverted because in the dialect of these speakers, the declarative version must have an auxiliary.

The majority of inconsistencies in question coding were in whether a question was prosodically marked or not. Furthermore, the knowledge that an utterance is intended as a question might cause a coder to perceive prosodic question-marking where none exists. We therefore collected the multi-word ID utterances from MOT1 file 11 (as described in the methods section) that did not contain noise in the recording and sorted them into declarative (42), *wh*-question (21) and *yes/no* question (36). These utterances were low-pass filtered at 400 Hz and presented to six naive adult raters, who were asked to judge whether they sounded more like questions or statements, on a scale of 1 (statement) to 7 (question). All six raters judged the *yes/no* questions to be more question-like than the declarative utterances in a non-paired t-test (Rater1: $t=3{\cdot}31$, $p=0{\cdot}001$; Rater2: $t=4{\cdot}57$, $p<0{\cdot}001$; Rater3: $t=7{\cdot}08$, $p<0{\cdot}001$; Rater4: $t=3{\cdot}84$, $p<0{\cdot}001$; Rater5: $t=5{\cdot}68$, $p<0{\cdot}001$; Rater6: $t=6{\cdot}22$, $p<0{\cdot}001$, two-tailed). However, only 1 out of 6 raters judged the *wh*-questions to be more question-like than the declarative utterances (Rater2: $t=1{\cdot}89$, $p<0{\cdot}05$, one-tailed), and the average judgments across all six raters were not significantly different.

We further categorized each utterance as 'falling', 'rising', 'U-shaped' or 'bell/sinusoidal', by inspection of the pitch contour, following Stern, Spieker & McKain (1982). Also according to their method, we excluded utterances with pitch contours smaller than 128 Hz. This left 60 utterances for analysis – 26 declarative utterances, 14 *wh*-questions, and 20 *yes/no* questions. Only 1 utterance (a *yes/no* question) was judged to be U-shaped, so this category was discarded. A chi-squared analysis found overall significant differences ($\chi^2(8, N=59)=34{\cdot}7$, $p<0{\cdot}001$) in the distribution of the remaining utterances. *Yes/no* questions were significantly more likely to have rising contours than chance ($\chi^2(2, N=19)=14{\cdot}3$, $p<0{\cdot}001$), with 76% of rising contours being *yes/no* questions. *Wh* questions were significantly more likely to have falling or bell/sinusoidal contours than chance ($\chi^2(2, N=14)=9{\cdot}1$, $p=0{\cdot}011$), with 29% of these contours being *wh*-questions. Declaratives were also shifted toward falling and bell/sinusoidal contours, at 55%, but this effect was not significant from chance ($\chi^2(2, N=26)=3{\cdot}09$, $p>0{\cdot}1$). While *yes/no* questions were clearly marked prosodically (consistent with the results of Stern *et al.*, 1982), our data suggest that *wh*-questions were not as easily distinguishable prosodically from declarative utterances.

The highly consistent presence of prosodic marking in *yes/no* questions suggests that prosodic information may be very useful for the infant in distinguishing the more complex interrogatives from the more straightforward SVO declaratives. While *wh*-questions were not prosodically distinguishable, they contained lexical markers of their question status. Such information might be helpful in focusing the infant's computational attention initially on the simpler forms and/or highlighting the variation between declarative and question forms.

Summary. Overall, we found that the speech to both infants contained a variety of utterances types, including a large number of questions and utterances of varying length and complexity. Although MOT1's speech contained more complex utterances than MOT2, both mothers produced significant numbers of paralinguistic utterances, single words, isolated phrases and full clauses. While isolated phrases were present in strong enough numbers to indicate a possible key role in acquisition, full clauses and multi-clause utterances were also evident in the speech of both mothers. This suggests a possible role for prosodic structure in providing information to infants about the syntactic structure of the language. Our transcriber coding also supported this idea, as a variety of prosodically isolated grammatical units were found within the utterances. The next section examines some prosodic characteristics of our corpus quantitatively.

Quantitative measures of prosodic/syntactic breaks

The results in the previous section suggest that prosodic information might play a relatively indirect role in the development of grammatical knowledge, by grouping sequences of words into smaller, syntactically-relevant units such as clauses. It is also possible that prosody provides infants with more direct access to information about syntactic boundaries at the level of phrases within larger syntactic units. While there is strong evidence that sentence/utterance boundaries are prosodically marked in typical speech to infants (e.g. Bernstein Ratner, 1986; Fisher & Tokura, 1996*a*), the evidence with respect to phrase boundaries is less clear. For one thing, prosodic marking of phrase boundaries is much less reliable. In some circumstances, for example when the subject is a pronoun, there is likely not to be any kind of prosodic boundary before the verb phrase (Gee & Grosjean, 1983). Furthermore, since ID speech tends to be very short (short sentences with pronominal subjects, or individual phrases and sentence fragments), there is much less opportunity for such cues to appear.

Perceptual experiments suggest that infants are able to detect phrases based on their prosodic characteristics in laboratory-controlled ID speech in some circumstances (e.g. Jusczyk *et al.*, 1992; Soderstrom *et al.*, 2003), but not when the availability of cues is more limited, as in the case of pronouns (Gerken, Jusczyk & Mandel, 1994). Fisher & Tokura's (1996*a*) analysis of speech to Japanese- and English-learning fourteen-month-olds found some cues to utterance-internal phrase boundaries, including segmental lengthening in English and intonational changes in Japanese. However, fourteen-month-olds have begun to build a lexicon, and potentially have access to much more sophisticated cues to syntactic boundaries than younger infants. Furthermore, the linguistic environment of fourteen-month-olds may be very different from that of younger infants, since fourteen-month-olds are more adept conversational partners. Because infants as young as six months may be using prosody to process speech, it is important to gain a better picture of the extent to which speech to younger infants contains phrase-level prosodic cues.

Utterance boundaries. We first compared utterance boundaries with all other boundaries in two separate MANOVAs. The first (raw score) comparison included the measures pre-boundary vowel duration, pre-boundary intensity, pause length and change in pitch from pre- to post-boundary syllable. The second (difference score) comparison included vowel duration difference, intensity difference, pause length and change in pitch. The methods section provides specific details on how these measures were obtained. Tables 7 and 8 summarize the statistical findings for all acoustical analyses. See Appendix 2 for descriptive statistics on these acoustic variables.

Utterance boundaries were highly distinguishable from other boundaries in both analyses for both mothers overall, and in the individual measures vowel duration, vowel duration difference and pause length. Average intensity was also significant for both mothers. Intensity difference and peak pitch change were non-significant. Because of the enormous differences found at utterance boundaries, these syllables were excluded from the remainder of the analyses.

Utterance-internal clause-level boundaries. We next examined whether utterance-internal clause-level boundaries differed from simple word boundaries. Previous analyses have often confounded clause and utterance boundaries. Given the highly significant effects at utterance boundaries compared with other syllable boundaries, we felt it important to determine whether non-utterance-boundary clause-level boundaries are identifiable on the basis of their acoustical properties. In this analysis we included both full clauses and isolated phrases or words within an utterance in this analysis. For example, for the utterance *Sit up, please, the right way.*, there were three clause-level units: *sit up*, *please* and *the right way*, and three simple word-level boundaries, after *sit*, *the* and *right*. Utterance-internal clause-level boundaries for both mothers were highly distinguishable from simple

TABLE 7. *Statistical data for MANOVAs of prosodic boundaries (raw scores)*

	Overall			Pause length			Vowel duration			Change in pitch			Intensity		
	F[a]	p[b]	η^2	F	p	η^2	F	p	η^2	F	p	η^2	F	p	η^2
Utterance boundaries															
MOT1	212·3 (4, 742)	**<0·001**	0·534	739·6 (1, 745)	**<0·001**	0·498	141·5 (1, 745)	**<0·001**	0·160	0·57 (1, 745)	>0·1	0·001	6·89 (1, 745)	**0·009**	0·009
MOT2	170·9 (4, 545)	**<0·001**	0·556	531·4 (1, 548)	**<0·001**	0·492	115·7 (1, 548)	**<0·001**	0·174	1·27 (1, 548)	>0·1	0·002	7·24 (1, 548)	**0·007**	0·003
Clause-level boundaries															
MOT1	17·58 (4, 336)	**<0·001**	0·173	35·97 (1, 339)	**<0·001**	0·096	31·62 (1, 339)	**<0·001**	0·085	11·83 (1, 339)	**0·001**	0·034	0·01 (1, 339)	>0·5	<0·001
MOT2	9·05 (4, 217)	**<0·001**	0·143	4·23 (1, 220)	**0·041**	0·019	26·31 (1, 220)	**<0·001**	0·107	1·56 (1, 220)	>0·1	0·007	4·92 (1, 220)	**0·028**	0·022
Question phrase boundaries															
MOT1	3·27 (4, 67)	**0·016**	0·163	2·11 (1, 70)	>0·1	0·029	2·01 (1, 70)	>0·1	0·028	6·62 (1, 70)	**0·012**	0·086	0·50 (1, 70)	>0·1	0·007
MOT2	14·4 (4, 63)	**<0·001**	0·478	2·14 (1, 66)	>0·1	0·031	7·98 (1, 66)	**0·006**	0·108	10·64 (1, 66)	**0·002**	0·139	21·00 (1, 66)	**<0·001**	0·241
Declarative phrase boundaries															
MOT1	0·454 (4, 126)	>0·5	0·014	0·086 (1, 129)	>0·5	0·001	0·240 (1, 129)	>0·5	0·002	1·67 (1, 129)	>0·1	0·013	0·018 (1, 129)	>0·5	<0·001
MOT2	0·297 (4, 119)	>0·5		0·729 (1, 122)	>0·1		0·314 (1, 122)	>0·5		0·001 (1, 122)	>0·5		0·089 (1, 122)	>0·5	

[a] Numbers in parentheses are degrees of freedom; the number above is the F-value.
[b] Significant p-values are marked in bold.

TABLE 8. *Statistical data for MANOVAs of prosodic boundaries (difference scores)*

	Overall			Pause			Vowel duration			Change in pitch			Intensity		
	F[a]	p[b]	η^2	F	p	η^2	F	p	η^2	F	p	η^2	F	p	η^2
Utterance boundaries															
MOT1	194·5 (4, 741)	**<0·001**	0·512	738·5 (1, 744)	**<0·001**	0·498	55·28 (1, 744)	**<0·001**	0·069	0·57 (1, 744)	>0·1	0·001	1·06 (1, 744)	>0·1	0·001
MOT2	167·8 (4, 543)	**<0·001**	0·553	641·1 (1, 546)	**<0·001**	0·540	25·82 (1, 546)	**<0·001**	0·045	1·02 (1, 546)	>0·1	0·002	1·96 (1, 546)	>0·1	0·004
Clause-level boundaries															
MOT1	19·30 (4, 336)	**<0·001**	0·175	35·97 (1, 339)	**<0·001**	0·096	28·73 (1, 339)	**<0·001**	0·078	11·83 (1, 339)	**0·001**	0·034	0·18 (1, 339)	>0·5	0·001
MOT2	9·82 (4, 217)	**<0·001**	0·153	4·23 (1, 220)	**0·041**	0·019	29·87 (1, 220)	**<0·001**	0·120	1·56 (1, 220)	>0·1	0·007	1·11 (1, 220)	>0·1	0·005
Question phrase boundaries															
MOT1	3·22 (4, 67)	**0·021**	0·161	2·11 (1, 70)	>0·1	0·029	2·10 (1, 70)	>0·1	0·029	6·62 (1, 70)	**0·012**	0·086	4·00 (1, 70)	**0·049**	0·054
MOT2	5·23 (4, 63)	**0·001**	0·249	2·14 (1, 66)	>0·1	0·031	1·11 (1, 66)	>0·1	0·017	10·64 (1, 66)	**0·002**	0·139	2·59 (1, 66)	>0·1	0·038
Declarative phrase boundaries															
MOT1	0·499 (4, 125)	>0·5	0·016	0·071 (1, 128)	>0·5	0·001	0·009 (1, 129)	>0·5	<0·001	1·67 (1, 129)	>0·1	0·013	0·349 (1, 129)	>0·5	0·003
MOT2	1·282 (4, 119)	>0·1	0·041	0·729 (1, 122)	>0·1	0·006	3·13 (1, 122)	0·079	0·025	0·001 (1, 122)	>0·5	<0·001	1·67 (1, 122)	>0·1	0·014

[a] Numbers in parentheses are degrees of freedom, the number above is the F-value.
[b] Significant p-values are marked in bold.

word boundaries using both raw and difference scores. For MOT1, all measures were individually significant except for average intensity and intensity difference. For MOT2, all measures were individually significant except peak pitch difference and intensity difference. These data suggest that non-utterance-boundary clause-level boundaries are acoustically distinguishable from simple word boundaries. Interestingly, the specific cues that contributed to the effect were not identical between the two mothers. While both mothers produced significant pausing and vowel duration cues, we found significant effects of peak pitch difference for MOT1 but not MOT2, and significant effects of average intensity for MOT2 but not MOT1. Whether these statistical differences indicate systematic differences in the way the mothers reflect clause-level boundaries is not clear, but they do suggest that it is important to consider clusters of cues in examining these prosodic effects.

Clause-internal phrase-level boundaries. The evidence for phrase-level effects is more complicated, in part because there are a number of different types of phrase boundaries one might consider, which are likely to have different prosodic characteristics. We therefore performed separate analyses on declarative and inverted question subject–predicate boundaries. For the declarative analysis, we specifically targeted the boundary between subjects and predicates (excluding auxiliaries) comparing it with the verb–argument boundary. This comparison was made because the subject–predicate and verb–argument boundaries were often adjacent in the utterance, and therefore this boundary provided a good control for effects of speech rate. A separate analysis comparing the subject–predicate boundary with all within-phrase word-level boundaries yielded similar results but is not reported here. For similar reasons, in the inverted question analysis, we compared the subject–predicate boundary with the auxiliary–subject boundary. We found no significant differences between phrase and word boundaries in the declarative analysis, and no interactions. An analysis specifically targeting multisyllable declarative subjects across MOT1's transcripts ($N=36$) failed to find evidence of prosodic cues, although the results of perceptual studies suggest that such cues exist in some non-pronoun declarative utterances (Jusczyk *et al.*, 1992; Soderstrom *et al.*, 2003), at least with laboratory-created stimuli.

However, in the inverted question analysis, both mothers showed significant differences in both raw and difference scores. For MOT1, there were individually significant differences in peak pitch difference and intensity difference, with a decrease in pre-phrase boundary pitch relative to the post-boundary syllable, and a decrease in relative intensity in the pre-phrasal boundary syllables, compared with pre-word boundary syllables. The remaining measures were non-significant. For MOT2, vowel duration, average intensity and peak pitch difference showed significant differences, with longer and louder vowels in pre-phrase boundary syllables, and lower pre-phrase boundary pitch relative to the post-boundary syllable. Pause length, vowel duration difference and intensity difference were non-significant.

Overall, subject–predicate boundaries in questions are acoustically distinguishable from simple word boundaries, but as with clause boundaries, these effects are manifest in different ways in different contexts. Clearly, the infant is not presented with a simple acoustical indicator of all subject–predicate boundaries. However, these boundaries (at least in inverted questions) do appear to be identifiable based on prosodic cues.

Summary. Our acoustic analyses were consistent with previous research in finding strong cues to utterances boundaries. Our findings went beyond previous work in demonstrating that clause-level boundaries WITHIN utterances were prosodically distinguishable from other syllables. Therefore infant-directed speech contains cues to clause-level boundaries over and above the cues to utterance boundaries. We also found evidence for prosodic information at the level of the phrase, but only in our analyses of interrogative utterances. We did not find evidence of prosodic cues to phrase boundaries in declarative utterances, even those with multisyllabic subjects. This finding therefore suggests a possible role for questions in providing infants with cues to verb phrase boundaries.

DISCUSSION

Consistent with previous studies, our analysis of the maternal input to two young infants found overall very simple utterances, short clauses with primarily pronoun subjects and little developmental change in complexity across the period of six to ten months. Consistent with Brent & Siskind (2001), we found a significant number of isolated words. Also consistent with previous work, but in contrast to the finding of van de Weijer (2002*a*) with a Dutch infant, we found a large number of questions.

Superficially, these findings present a picture of the input in which there is little role or necessity for prosodic information. If infants are hearing primarily individual words and phrases, or short clauses with pronoun subjects, there is little necessity for infants to attend to prosodic structure beyond the highly salient utterance boundaries – indeed such prosodic structure would not be present for the infants to find. However, our analyses suggest two things to contradict this view. First, while utterances tend to be short, they also tend to contain more than one syntactic unit, whether two words that do not combine syntactically into a phrase or a phrase or clause preceded by, for example, a phatic word. Second, both mothers (especially MOT1) produced some utterances with sufficient length and complexity to benefit from (and contain) internal prosodic organization, such as multi-clause utterances and *yes/no* questions.

In our qualitative coding, we noted prosodic breaks that isolated groups of words from the main utterance. This prosodic information served to isolate or highlight syntactic units within utterances at a number of different levels of analysis, from clauses to phrases to individual words. This impression was supported by our quantitative analysis, in which we found evidence for the presence of utterance-internal cues to clause-level boundaries, as well as clause-internal phrase-level cues to a major syntactic boundary (the subject–VP boundary). These phrase-level cues appeared to be exclusively present in *yes-no* questions, suggesting that the special prosodic properties of *yes-no* questions might play an important role in the infant's initial organization of linguistic information. It is interesting that it is just this prosodic distinction between *yes-no* questions and declaratives that was highlighted by adult judgments of our speech samples.

The findings of this study present a complex picture of the prosodic properties of maternal speech and the roles it might play in early grammatical development. While the possibility of prosodic bootstrapping is sometimes presented as an either/or question, prosodic information may play a role in language acquisition in different ways at different times (and potentially also for different infants). We will briefly examine some of these possibilities.

Isolating utterances in noisy environments

While MOT2's recordings were generally made in a quiet environment, the environment of MOT1's recordings produced a large amount of overlapping speech. These utterances suggest an interesting component of the input to at least some infants – utterances that overlap in time. For INF1, this chaotic speech environment was probably the rule rather than the exception. For infants in such environments, the familiarity of the maternal voice might allow infants to selectively attend to one stream of speech (Barker & Newman, 2004). However, this does not explain how infants are able to separate the different streams from each other acoustically in order to detect the familiar voice. Prosodic information, by providing an acoustic envelope to utterances, could help infants like INF1 to isolate utterances not only temporally, but also spatially.

'*Cleaning up*' *speech*

Early in language development, infants are presented with the task of determining what constitutes the meaningful components of the speech input they are hearing. This speech input includes vocal play, gasps and sighs and other speech-like vocalizations. While these expressions might play an important role in communicating affect and engaging the infant, they are not directly linguistic expressions per se. It would be very useful for infants to distinguish common paralinguistic expressions like *oh* and *aah* from important linguistic functional elements like *the* and *a*. Fully 20–40 percent of the major prosodic breaks transcribed in our analysis separated elements of this kind from the rest of the utterance. It is worth exploring further the extent to which acoustic measures support our coders' impressions of these prosodic breaks. Prosody might play an important role in separating these paralinguistic elements from the rest of the utterance and highlighting the difference between these elements and real linguistic elements like determiners (compare *oh, baby!* to *a baby!*).

Finding words

Discovering the boundaries between words in fluent speech is a crucial component of breaking into language. Not only are words an integral part of discovering the grammatical structure of the language (and of course the ability to actually interpret it, which is the goal of language development), but finding individual words in speech may help the infant to discover additional words (Brent & Siskind, 2001; Bortfeld, Morgan, Golinkoff & Rathbun, 2005). Close to 40 percent of the elements separated from the main utterance by prosodic information in our analysis were single words. This suggests that prosodic information might substantially facilitate the process of early lexical development. It is interesting to note that a great number of these individual words were names, particularly the name of the infant and pet names for the child (e.g. *cutie*, *handsome* or *sweetie*), but also names for the mother, other adults and children present or labels for objects. It is little wonder that familiar names can be used for top-down segmentation of adjoining novel words as early as six months (Bortfeld *et al.*, 2005).

Finding phrases and clauses

Our analysis also found a significant number of isolated phrases and clauses generated by prosodic grouping in speech to infants as young as six months. Qualitative transcriber codings were supported by the acoustic analyses, which found strong cues to utterance-internal clause-level boundaries. This, together with the behavioral findings that infants use prosodic information to group words (e.g. Soderstrom *et al.*, 2003) supports the notion that even young infants might usefully rely on a strategy of 'divide and conquer' (Jusczyk, 1998), where larger utterances are grouped together into syntactically relevant prosodic units for further analysis. It is worth noting that while individual phrases were found both in the analysis of isolated utterances and prosodically grouped units, full clauses were much more common, both for MOT1 and MOT2. However, the vast majority of individual clauses contained pronoun subjects. As Fisher & Tokura (1996*b*) point out, these acoustically reduced subjects might allow infants to detect additional phrase-level syntactic units, particularly verb phrases.

Finding syntactic boundaries directly

The more radical possibility to consider is that infants might obtain more direct information about phrase-level syntactic boundaries based on acoustical properties of the input. Instead of (or as well as) finding a verb phrase within a clause based on having previously heard that verb phrase as an isolated (or semi-isolated) utterance, infants might be able to detect the presence of subject–predicate boundaries by noticing specific acoustical properties of the adjacent syllables. There are several arguments against this idea, some of which have been previously discussed:

- Phrase-level prosodic boundary information is unreliable: subject–predicate boundaries are not reliably marked in the input.
- Prosody is dictated by aspects of the language other than syntax: it has long been understood that syntax is not the ONLY influence on the prosodic structure of an utterance.
- ID utterances are short: with an average MLU of 4, there is little room for prosodic breaks to occur.
- ID subject NPs consist primarily of pronouns: pronoun subjects are highly reduced, and unlikely to contain prosodic boundary information.

However, it may be premature to dismiss the role of direct evidence for prosodic information in early grammatical development. Whether the correlation between prosodic structure and syntax is consistent enough to be useful to infants could be explored through additional statistical analysis and/or computational models. Certainly the behavioral evidence suggests that infants pay attention to phrase-level prosodic information. It may be that this information is reliable in some contexts (such as inverted questions), but not others (such as pronoun declaratives).

Even in cases where the input prosody does NOT coincide with syntax, prosodic information may nevertheless play a role in the infant's analysis of linguistic structure. It does not matter so much whether prosody faithfully reflects the intended syntactic structure of input utterances as whether infants expect it to do so. Thus, if infants do rely on phrase-level prosodic information in forming structural hypotheses, systematic prosody–syntax mismatches might lead to systematic reanalyses of syntactic structure. For example, as noted, pronominal subjects of declarative sentences are prosodically packaged with the verb phrase. If infants are guided by prosody in their interpretation of syntactic grouping of words, this may lead them to infer that nominative pronouns must co-occur with a verb phrase. Other subject NPs, however, which are not prosodically packaged with the VP would not be subject to this stricture. In this light, it is interesting to note that the distribution of nominative pronouns has become much more restricted than that of other NPs in English. Thus, whereas (1a) was at one time the preferred form, it now is considered stilted and hypercorrect; (1b) in which an objective pronoun is used, is more acceptable, and (1c), where the nominative pronoun is accompanied by a tense-bearing element, is also possible. Moreover, whereas either (2a) or (2b) and (2d) are acceptable, (2c) clearly is not.

(1a) ?Madison is taller than he.
(1b) Madison is taller than him.
(1c) Madison is taller than he is.
(2) Who did it?
  a. John.
  b. John did.
  c. *He.
  d. He did.

Our findings, together with previous research on the complexity of ID speech (e.g. Snow, 1977) suggest that even speech to infants under one year contains some utterances long enough to contain phrase-level prosodic cues – primarily in MOT1's speech, but also to a smaller extent in that of MOT2. Furthermore, the presence of others in the environment, both adults and children, may have a strong influence on the length of maternal utterances. Our results with MOT1 are suggestive of this idea. Additionally, van de Weijer's (2002*a*) work suggests that exclusively ID speech may be a minority of the language input to an infant. While infants show a preference for listening to ID speech, this does not mean that infants only pay attention to speech directed at them. It may be that AD speech plays a greater role in early language acquisition than is typically considered.

Our acoustic analyses, like those of Fisher & Tokura (1996*a*) with older infants, found evidence that the acoustical properties of the boundaries between subjects and predicates differ from those of phrase-internal word-level boundaries. If we consider these acoustical properties across all sentence types, the cues appear weak and unreliable. However, when we examined declaratives separately from inverted questions, which was not done in the original Fisher & Tokura study, the declarative sentences showed no obvious cues to these phrase-level boundaries, while the inverted questions showed significant effects. This suggests that these two sentence types may play different roles as sources of information about grammatical structure in English.

Ultimately, acquisition is driven by the RELATIONSHIP between infants' perceptual biases and the properties of speech input. The exploration of properties of infants' speech input must be guided by the nature of infants' perceptual abilities and biases. The reverse is also true – experiments in infant speech perception must be informed by our understanding of the natural speech input to the infant. Detailed analyses of infant-directed speech of the sort we have presented here are crucial for promoting such understanding.

REFERENCES

Barker, B. A. & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition* **94**, B45–B53.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother–child speech. *Phonetics* **14**, 303–309.

Blauww, E. (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication* **14**, 359–75.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International* **5**, 341–45.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science* **16**, 298–304.

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

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P. *et al*. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual.* San Diego: Singular Publishing Group.

Fernald, A. & McRoberts, G. (1996). Prosodic bootstrapping: A critical analysis of the argument and the evidence. In J. L. Morgan & K. Demuth (eds), *From signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 365–88. Hillsdale, NJ: Erlbaum.

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.

Fisher, C. & Tokura, H. (1996*a*). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development* **67**, 3192–218.

Fisher, C. & Tokura, H. (1996*b*). Prosody in speech to infants: Direct and indirect cues to syntactic structure. In J. L. Morgan & K. Demuth (eds), *Signal to syntax*, 343–64. Mahwah, NJ: Erlbaum.

Furrow, D., Nelson, K. & Benedict, H. (1979). Mothers' speech to children and syntactic development: some simple relationships. *Journal of Child Language* **6**, 423–42.

Gee, J. P. & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* **15**, 411–25.

Gerken, L. A., Jusczyk, P. W. & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: Nine-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition* **51**, 237–65.

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*, 3–48. Cambridge: Cambridge University Press.

Jusczyk, P. W. (1998). Constraining the search for structure in the input. *Lingua* **106**, 197–218.

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**, 252–93.

Kaye, K. (1980). Why we don't talk 'baby talk' to babies. *Journal of Child Language* **7**, 489–507.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* **3**, 129–40.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*, 3rd 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**, 155–80.

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.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* **38**, 1–135.

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*, 109–149. Cambridge: Cambridge University Press.

Papousek, M., Papousek, H. & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. *Journal of Psycholinguistic Research* **16**, 491–516.

Penman, R., Cross, T., Milgrom-Friedman, J. & Meares, R. (1983). Mothers' speech to prelingual infants: a pragmatic analysis. *Journal of Child Language* **10**, 17–34.

Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development* **44**, 182–85.

Pinker, S. (1994). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

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

Shady, M. (1996). Infants' sensitivity to function morphemes. Unpublished doctoral dissertation, State University of New York, Buffalo.

Sherrod, K. B., Friedman, S., Crawley, S., Drake, D. & Devieux, J. (1977). Maternal language to prelinguistic infants: syntactic aspects. *Child Development* **48**, 1662–65.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language* **4**, 1–22.

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**, 249–67.

Soderstrom, M., White, K. S., Conwell, E. & Morgan, J. L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy* **12**, 1–29.

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

Stern, D. N., Spieker, S. & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology* **18**, 727–35.

van de Weijer, J. (2002*a*). How much does an infant hear in a day? Paper presented at the GALA 2001 Conference on Language Acquisition.

van de Weijer, J. (2002*b*). Terminal rises in infant-directed and adult-directed question. In *Speech, Music and Hearing, Quarterly Progress and Status Report, Vol. 44* (*Proceedings of Fonetik, 2002*), 5–8, Stockholm, Sweden.

Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992). Segmentation durations in the vicinity of prosodic boundaries. *Journal of the Acoustical Society of America* **91**, 1707–717.

APPENDIX 1: SYNTACTIC AND PROSODIC TRANSCRIPTION CODES

Category	Description	Examples
Nonsense	Vocal sounds with no semantic content.	Raspberries, babbling
Quasi-word	Carries some semantic content, but does not fulfill the requirements of being a fully lexical word.	Voluntary laughter, gasps, sound effects, animal noises, exclamations and interjections (*oh* and *hm*, etc.), phonemically incomplete words due to disfluency
Word: *Isolated*	One and only one full lexical word. A subset of the Word: All category.	No!
Word: *All*	At least one full lexical word. May also contain quasi-words or nonsense, or multiple lexical words separated by a prosodic break that do not combine phrasally.	Yes, please. Oh, yeah? No!
Phrase	Contains more than one lexical word not separated by major prosodic break, but does not contain both a subject and a predicate. Includes imperatives and elided utterances.	Little fat+o? agagagaga, Mister Chatterbox. Sit up, please, the right way. To you too!
Clause	Includes a single subject and predicate.	What are you telling me? You shouldn't (h)ave hit your head into that tray.
Multiclause	Contains multiple clauses – excluding repetitions due to disfluency.	Let me see you clap. You just want the wire, that's what you think you're gonna get.
Fluent break	Fluent, syntactically consistent, prosodic break, usually marked by a comma in the transcript.	Oh, you think you need to touch it. But no pulling, please. I know, they need to go down to eat.
Disfluent	Caused by an obvious disfluency, retracing or repetition.	I don't know why anyone [//] nobody has any clean pajamas in this house.
Other	Not a disfluency with restart, but sounds syntactically or prosodically 'odd'. Usually a pause for thought.	I just think that's # in too close. Without having a # twenty minute pot of coffee.

APPENDIX 2: MEANS, Ns AND STANDARD DEVIATIONS FOR PROSODIC BOUNDARIES

	Pause (ms)		Vowel duration (ms)		Vowel duration difference (ms)		Change in pitch (semitones)		Intensity (dB)		Intensity difference (dB)	
Utterance vs. non-utterance	U	N	U	N	U	N	U	N	U	N	U	N
MOT1	1620[a] (1772) N:92	6 (25) N:709	209 (134) N:92	103 (65) N:711	85 (178) N:61	−10 (85) N:705	−0·545 (9·46) N:59	0·139 (6·39) N:688	70·9 (4·83) N:92	73·7 (7·31) N:711	0·682 (10·4) N:61	0·060 (8·8) N:705
MOT2	2138 (2496) N:124	16 (58) N:540	246 (213) N:125	124 (77) N:541	81 (271) N:62	23 (119) N:538	−1·86 (9·60) N:55	−0·657 (7·16) N:497	65·1 (6·62) N:125	66·5 (7·11) N:542	−1·39 (9·18) N:63	−0·173 (7·7) N:543
Clause-level vs. word-level	C	W	C	W	C	W	C	W	C	W	C	W
MOT1	26 (49) N:58	4 (16) N:293	144 (83) N:58	98 (61) N:295	47 (93) N:58	−24 (92) N:291	3·03 (7·62) N:56	0·078 (5·47) N:285	73·8 (4·39) N:58	73·8 (8·35) N:295	0·616 (3·26) N:58	0·315 (8·24) N:292
MOT2	60 (107) N:54	14 (58) N:198	166 (87) N:54	116 (79) N:200	49 (104) N:54	−46 (125) N:199	1·21 (8·93) N:47	−0·220 (6·64) N:177	68·0 (6·90) N:54	65 (7·54) N:200	1·22 (5·67) N:54	−0·855 (7·87) N:199
Subject–Predicate vs. Auxiliary–Subject (questions)	SP	AS	SP	AS	SP	AS	SP	AS	SP	AS	SP	AS
MOT1	3 (12) N:35	0 (0) N:38	104 (119) N:35	75 (42) N:38	4 (116) N:35	−36 (121) N:37	−0·73 (3·97) N:36	1·92 (4·71) N:37	73·0 (3·92) N:35	72·2 (4·23) N:38	−1·94 (3·50) N:35	−0·273 (3·59) N:37
MOT2	10 (30) N:35	1 (7) N:38	124 (45) N:35	89 (51) N:38	−52 (97) N:35	−36 (65) N:38	−3·80 (5·36) N:34	0·351 (5·14) N:34	66·6 (6·37) N:35	60·0 (6·39) N:38	−2·45 (4·41) N:35	−4·98 (9·40) N:38

APPENDIX 2 (*Cont.*)

	Pause (ms)		Vowel duration (ms)		Vowel duration difference (ms)		Change in pitch (semitones)		Intensity (dB)		Intensity difference (dB)	
Subject–Predicate vs. Verb–Argument (declaratives)	SP	VA	SP	VA	SP	VA	SP	VA	SP	VA	SP	VA
MOT1	3 (11) N:48	5 (18) N:87	96 (47) N:48	91 (38) N:87	−11 (82) N:47	−13 (52) N:87	−1·07 (5·89) N:46	0·391 (6·33) N:85	73·0 (5·10) N:48	73·3 (10·9) N:87	−1·71 (4·42) N:46	−0·309 (11·00) N:87
MOT2	3 (13) N:36	7 (30) N:93	125 (58) N:36	131 (61) N:93	−20 (98) N:36	9 (77) N:92	−1·36 (8·79) N:36	−1·41 (7·84) N:88	67·7 (5·91) N:36	67·8 (6·10) N:93	−0·143 (5·60) N:36	1·43 (5·13) N:92

[a] First value is the mean. Standard deviation is given in parentheses.