Running Head: PROSODIC BOOTSTRAPPING OF PHRASES

The Prosodic Bootstrapping of Phrases: Evidence from Prelinguistic Infants

Melanie Soderstrom[1]

Department of Psychological and Brain Sciences

Johns Hopkins University

Amanda Seidl

Departments of Psychological and Brain Sciences & Cognitive Science

Johns Hopkins University

Deborah G. Kemler Nelson

Department of Psychology

Swarthmore College

Peter W. Jusczyk

Departments of Psychological and Brain Sciences & Cognitive Science

Johns Hopkins University

Corresponding Author:
Melanie Soderstrom
Tel: 401-863-3959
Fax: 401-863-2255
Brown University
Dept. of Cognitive and Linguistic Sciences
100 Thayer St. /Box 1978
Providence RI 02912-1978
melsod@brown.edu

Abstract

The current study explores infants' use of prosodic cues coincident with phrases in processing fluent speech. After familiarization with two versions of the same word sequence, both 6- and 9-month-olds showed a preference for a passage containing the sequence as a noun phrase over a passage with the same sequence as a syntactic non-unit. However, this result was found only in one of two groups, the group exposed to a stronger prosodic difference between the syntactic and non-syntactic sequences. Six-month-olds were tested in the same way on passages containing verb phrases. In this case, both groups preferred the passage with the verb phrase to the passage with the same word sequence as a syntactic non-unit. These results provide the first evidence that infants as young as 6 months old are sensitive to prosodic markers of syntactic units smaller than the clause, and, in addition, that they use this sensitivity to recognize phrasal units, both noun and verb phrases, in fluent speech. This ability to use phrase-level prosodic cues is variable, however, and appears to depend on the strength or number of cues associated with these syntactic units.

Key Words: prosody, prosodic bootstrapping, phrase segmentation, speech segmentation, language acquisition, infant speech perception

The Prosodic Bootstrapping of Phrases: Evidence from Prelinguistic Infants

An important aspect of acquiring a natural language is discovering the relevant syntactic and prosodic units in the input. Consider one important unit of speech, the Prosodic Word. Considerable work has focused on how and when infants find the boundaries between individual words in speech (e.g., Brent & Cartwright, 1996; Johnson & Jusczyk, 2001; Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999; Myers et al., 1996; Saffran, Aslin, & Newport, 1996). Findings suggest that infants are able to segment trochaic (strong-weak) words from fluent speech by 7.5 months (Jusczyk et al., 1999).

Strings of speech also contain prosodic information about syntactic structure (e.g., Fisher & Tokura, 1996a; Gleitman & Wanner, 1982). Infants may use this information to learn something about the grammatical structure of their language. Gleitman and Wanner (1982) originally proposed that acoustic cues in the speech stream might provide infants with cues to syntactic boundaries, even before lexical knowledge is available (see also Gleitman, Gleitman, Landau, & Wanner, 1988; Morgan & Newport, 1981; Peters, 1983; Pinker, 1984). This proposal has become known as “prosodic bootstrapping”, or “bootstrapping from the signal”. The hypothesis presupposes three elements: First, that syntax is reliably correlated with acoustic properties; second, that infants are sensitive to these properties; and third, that infants use these acoustic cues in their processing of speech.

What acoustic cues to syntactic boundaries are present in the speech stream that may be noticed and subsequently used by infants? A few cues to the location of boundaries in the speech stream have been examined over the years. One of these cues is preboundary lengthening, or lengthening of the rhyme, the part of a syllable that does not include the initial consonant/s – for example, [æt] in 'cat' or [u] in 'Lou', at the end of a grammatical unit (Beckman & Edwards, 1990, 1992; Cooper & Paccia-Cooper, 1980; Klatt, 1975; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). Another cue is pause duration. For example, speakers tend to produce longer pauses at word boundaries when they coincide with clause boundaries (Cooper & Paccia-Cooper, 1980; Scott, 1982). Change in f_0 may also be a cue to a syntactic boundary. Specifically, a special kind of intonation pattern or tone pattern may occur at a phrase or clause boundary, marking it as such (Beckman & Pierrehumbert, 1986). For example, in many languages the high-rise at the end of a yes/no question marks a sentential/clause boundary of a particular type. Although syntactic units are not always reliably marked in the speech stream (Beckman & Edwards, 1990; Gerken, Jusczyk, & Mandel, 1994), sufficient cues may be present in speech to be potentially useful to infants (Jusczyk, 1998a).

There is considerable evidence that infants *are* sensitive to prosodic cues in the speech stream. For example, infants are capable of discriminating acoustic properties such as pitch change by 1-2 months old (Kuhl & Miller, 1982; Morse, 1972). By 4.5 months, infants begin to show sensitivities to certain prosodic markers in fluent speech, preferring passages with artificial pauses inserted at clause boundaries rather than other places in the sentence (Juscyzk, Hohne, & Mandel, 1995; see also Hirsh-Pasek et al., 1987; Kemler Nelson et al., 1995; Morgan, Swingley, & Miritai, 1993). Prosodic cues also influence infants' memory for a series of words. Mandel and colleagues (Mandel, Jusczyk, & Kemler Nelson, 1994; Mandel, Kemler Nelson, & Juscyzk, 1996) found that

2-month-olds dishabituated to a greater degree to a change in word order or a minimal phonetic change when words were linked prosodically as in a clause rather than spoken as either a simple list of words or two contiguous, incomplete clause fragments. More recent work has found that by 6 months, infants are able to use prosodic information coincident with clausal units in their processing of fluent speech (Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000).

These findings make a strong case for infants' use of prosodic cues in their processing of clauses. With respect to the use of prosodic cues that mark phrases, the story is more complex. Syntactically, clauses constitute the largest unit of constituent structure and meaning, and are often produced in isolation by one participant in a conversation (e.g., Did you walk the dog?). Phrases (for example, noun phrases, such as "the tall man", or verb phrases, such as "walks his dog") are sub-components of clauses, and provide the building blocks of clauses. The prosodic characteristics associated with clause boundaries in infant-directed speech are strong and fairly reliable. By contrast, the prosodic cues associated with phrases in infant-directed speech are weaker and may be unreliable (Beckman & Edwards, 1990; Fisher & Tokura, 1996a, 1996b; Gerken et al., 1994).

Yet phrases are crucial to language acquisition, and the ability to parse the speech stream at the level of the phrase would provide language learners with a window into the internal structure of sentences that would be unavailable if infants were only capable of segmenting fluent speech into clauses. Since syntax is, in an important way, "about" phrases and their internal structure, it is important to determine how infants might unpack clausal units in speech in order to discover these finer units embedded within them. For instance, one task that requires intra-clause analysis is determining where in the sentence the subject appears. Across most languages, Agents tend to appear as subjects. This fact has led several researchers (e.g. Pinker, 1989) to argue that learners could determine the position of the subject in their native language by identifying the position of the noun phrase that labels the Agent. In order to do this, however, infants would be aided considerably by being able to segment noun phrases from within clauses. Furthermore, the addition of prosodic cues has been shown to help adults learn the phrase structure of an artificial language (Morgan, Meier, & Newport, 1987), suggesting that infants might be similarly aided. The idea that phrase-level prosodic cues may be unreliable, however, potentially undermines these ideas.

Fisher and Tokura (1996b) proposed instead that infants are able to glean a fair amount of information regarding phrasal units without using prosody itself, from elliptical speech and other kinds of sentence fragments found in fluent infant-directed speech, as in (1), and imperatives, as in (2).

(1) What are you doing? *Washing dishes*.

(2) *Put that down*!

Another clause-level source of information for infants about phrasal structure might come from hearing different clauses produced in succession, and making comparisons across the words and/or word sequences they have in common (Gerken, 1996; Gerken et al., 1994; Jusczyk & Kemler Nelson, 1996).

An alternative idea, outlined in Jusczyk (1998a), is that infants *do* use prosodic cues that mark phrases in their segmentation of speech, even though such cues are not fully reliable. Jusczyk argued that infants employ a “divide and conquer” strategy, using prosody to parse speech into smaller units for analysis. Even though not all units that are recovered constitute well-formed syntactic units, prosodic cues are reliable often enough to allow infants to make progress in the acquisition of syntax. The possibilities outlined by Fisher and Tokura (1996b), Gerken (1996), and Jusczyk (1998a) are not necessarily mutually exclusive. It is likely that infants use a number of sources of evidence to learn the location of phrases in the target grammar.

Just how different are the prosodic features of phrases and clauses? Analyzing the acoustic properties of experimentally controlled speech by adults, Cooper and Paccia-Cooper (1980) found that the strength of prosodic cues varied with the hierarchical structure of grammar, with higher boundaries containing stronger cues. Across studies, prosodic constituents associated with a clause (usually corresponding to an Intonational Phrase[2]) are marked with larger and longer articulations than are those associated with a syntactic phrase (often corresponding to a Phonological Phrase), both with respect to cues such as pause duration, preboundary lengthening, and intonation, as well as with other cues in the speech stream. For example, Cho and Keating (1999) found that consonants at the beginnings of Intonational Phrases have both larger and longer linguopalatal contact (contact between the tongue and the top of the mouth) than consonants at the beginning of Phonological Phrases. Shattuck-Hufnagel and Turk (1996) review several similar findings. In addition, there are other acoustic phenomena that occur to a greater degree in Intonational Phrases. For example, Dilley, Shattuck-Hufnagel & Ostendorf (1996) found that for words that may be glottalized, glottalization occurs more often at the beginning of an Intonational Phrase than at the beginning of an Intermediate Phrase or Phonological Phrase.

In a first effort to evaluate infants' sensitivity to the prosodic cues that correlate with phrase boundaries, Jusczyk et al. (1992) employed the pause-insertion paradigm, previously used to examine infants' sensitivity to prosodic cues to clausal units (Hirsh-Pasek, et al., 1987). Rather than inserting pauses at clause boundaries, they inserted pauses at major phrase boundaries – namely, the boundary between the subject noun phrase and the verb phrase (VP). This will be referred to as the subject/VP boundary in the remainder of the paper. Jusczyk et al. found that 9-month-old infants showed a preference for passages with pauses coincident with this boundary over passages where the pauses were inserted elsewhere in the sentence. As with the clause studies, this was interpreted as evidence that infants are sensitive to this syntactic boundary based on the associated prosodic cues. However, Jusczyk et al. did not find a preference for the coincident passages with 6-month-olds.

Further insights about infants' sensitivity to phrase-level prosodic cues (i.e., cues associated with phrase boundaries) was provided in a follow-up study which examined different sentence types (Gerken et al., 1994). Gerken et al. compared sentences such as (3) with sentences such as (4). The more frequent prosodic boundaries are indicated by parentheses.

(3) (Joe) (kissed the dog).

(4) (He kissed) (the dog).

In sentences of the type exemplified in (3), speakers are likely to produce a prosodic boundary before the verb "kissed", which coincides with the subject/VP syntactic boundary. However, in sentences of the type exemplified in (4), which contains a weak pronoun, speakers either do not produce a salient prosodic boundary, or place the prosodic boundary after the verb “kissed” (e.g., Gee & Grosjean, 1983). Nine-month-old infants showed a preference for passages where the pause was located before the verb when sentences such as (3) were used as stimuli, but showed no preference for placement of the pause either before or after the verb when sentences with pronoun subjects such as (4) were used. In other words, infants demonstrated sensitivity to the syntactic boundary only in the first case, when it more reliably coincided with a prosodic boundary in natural speech.

The current experiments constitute a first attempt to garner evidence that infants not only are sometimes sensitive to the prosodic cues marking phrasal units, but that they deploy this sensitivity in their processing of fluent speech. The results of Jusczyk et al. (1992) and Gerken et al. (1994) demonstrate that, under some circumstances, infants are sensitive to the acoustic properties of phrase boundaries, but there is no evidence yet that infants actually use this sensitivity in the processing of the phonological information that is bracketed by these cues. Does infants’ sensitivity to prosodic cues to phrasal units have psychological consequences? Does it affect how infants encode and remember fluent speech?

Evidence that the prosodic information marking clauses has such effects was presented by Nazzi et al. (2000). Nazzi et al. found that infants were better able to recognize a word sequence embedded in continuous speech when the sequence constituted a prosodically well-formed clausal unit rather than a version of the same word sequence formed from the last words of one clause and the beginning words of the next. After familiarization with a word sequence spoken both as a well-formed unit (a clause) and as a syntactic non-unit (a word sequence that crossed a clause boundary), infants showed a preference for a passage containing the well-formed syntactic unit over a passage containing the non-unit. Nazzi et al. proposed that prosodic well-formedness facilitated the bracketing of the clausal unit in continuous speech. Hence, infants were better able to detect the recurrence of the sequence when it constituted a prosodically well-formed unit.

If infants use prosodic cues coincident with *phrases* in their processing of fluent speech, then the same principle should hold. After familiarization with a word sequence in both forms, infants should prefer a passage containing the sequence when it constitutes a prosodically well-formed phrasal unit than when it constitutes a word sequence that spans a major phrase boundary. The experiments reported in this paper make such assessments, using both subject noun phrases (Experiments 1 and 2) and verb phrases (Experiment 3) as the well-formed phrasal units.

Experiment 1

Following Jusczyk et al.'s (1992) success in demonstrating sensitivity to cues marking major phrase boundaries in 9-month-olds, Experiment 1 explores whether 9-month-olds actually deploy this sensitivity to cues internal to the clause in their processing of fluent speech. This effort is encouraged by Nazzi et al.'s (2000) suggestion that even 6-month-olds make use of prosodic cues correlated with clauses in speech processing. Still, there are reasons to question whether a parallel effect will obtain at the level of the phrase, both because prosodic cues to phrases are less reliable than those to clauses (as previously discussed), and because the relevant prosodic markers themselves may be different for phrases and clauses. For example, pause lengthening at the boundary (which was found in the stimuli employed by Nazzi et al.) may be a less pervasive characteristic of phrase boundaries than clause boundaries.

Method

Participants. Thirty-two infants (11 males, 21 females) from monolingual English-speaking households in the Baltimore area were tested. The infants had an average age of 39 weeks and 0 days (range = 37 weeks; 0 days to 41 weeks; 0 days) or approximately 9 months. Nine additional infants were tested but not included in the study. Three infants were not included for the following reasons: crying or fussing (2), average listening time less than 3 s during the test phase (1). Another 6 infants were excluded because they did not meet a listening criterion of more than 7 s on two separate trials for each stimulus type during the test phase. This criterion insured that infants would hear each familiarized sequence played to completion at least two times during the test phase.

Stimuli. The stimuli were read by a young adult female native speaker of American English from Westchester County, New York. They were recorded in a sound-proofed booth. The speaker was blind to the specific purpose of the experiment, and was given no particular instructions about how to produce the stimuli, except to produce them in natural, happy-sounding, infant-directed speech.[3] The following passages were recorded:

(I) At the discount store, **new watches for men** are simple and stylish. In fact, some **people # buy the whole** supply of them.

(II) In the field, the old frightened **gnu # watches for men** and women seeking trophies. Today, **people by the hole** seem scary.

Each passage contained two target sequences: a well-formed syntactic unit, a noun phrase (NP), and a syntactic non-unit (NU) that spanned a subject/VP boundary. The target sequences are in boldface print, and a hash mark (#) has been inserted to indicate the subject/VP boundary in the non-unit version of each sequence. Neither the hash mark nor the boldface print was present in the stimuli read by the speaker. These sentences were constructed with relatively long (5-syllable) subject noun phrases in order to increase the likelihood of prosodic cues at the subject/VP boundary (Gerken et al., 1994). Both passages contained the same two phonological word sequences. However, the sequence that was a well-formed syntactic unit, a noun phrase (NP), in Passage I was the non-unit (NU) in Passage II, and vice versa. Specifically, in Passage I, "**new watches for men**" constituted a well-formed noun phrase (NP), whereas **"people # buy the whole"** was a syntactic non-unit (NU). By contrast, in Passage II, "**people by the hole"** constituted a well-formed noun phrase (NP), while "**gnu # watches for men**" was a non-unit (NU).

A note about the ecological validity of these stimuli: These sentences might appear deceptively complex because of the sophistication of the vocabulary. The somewhat awkward vocabulary is a consequence of the design requirement to match the word sequences in the two passages phonologically. Crucially, there is nothing particularly odd about the phonotactics of the words, and, since the issue at hand is the use of prosodic information to extract structure, not meaning, we presume no familiarity with the words themselves, especially for these prelinguistic infants. In addition, syntactically, these stimuli are reasonably simple. Each sentence is composed of three phrases: a locative prepositional phrase (separated by a comma in the text), a subject noun phrase, and a verb phrase. The prepositional phrase was included, and placed at the front of the sentence, so that the beginning of the subject NP would not coincide with a clause boundary. This was done in order to afford the clearest contrast between Nazzi et al.'s (2000) test for the impact of clause-marking cues and the current test for the impact of phrase-marking cues.

Another concern might be the length of the phrases making up these sentences. The noun phrases are relatively long — 5 syllables, and thus it is possible that the acoustic properties of our stimuli may be limited to those sentences containing subject noun phrases this long or longer. But it is not unreasonable to expect that infants in the early stages of language acquisition sometimes hear noun phrases this long. An example produced by the mother in the first Eve file from the Brown corpus (Brown, 1973; CHILDES database, MacWhinney, 2000) is "Oh, dolly's other shoe is on her foot" (Eve 01), and other such examples exist. Therefore, we believe that our stimuli afford a reasonable first test of infants' sensitivity to phrase-level prosody, at least with respect to sentences with relatively long subject noun phrases.

The familiarization stimuli consisted of the four boldface target sequences, extracted from passages (I) and (II). The sequences were each approximately 1 second long. Silence was added to the end of each sequence up to 2.0 s. The sequence was repeated 8 times, to make a final sound file of 16.0 s for each sequence.

The test stimuli consisted of the two recorded passages. The durations of these passages were 7.8 s and 7.9 s. Silence was added to the end of each of these passages up to 8.5 s. Each passage was repeated 3 times, to make a final sound file of 25.5 s for each test passage.

Measurements were made of three of the acoustic properties/cues that have been shown to coincide with syntactic constituent boundaries in continuous speech. These were: preboundary lengthening (e.g., Beckman and & Edwards, 1990; Cooper & Paccia-Cooper, 1980; Klatt, 1975; Wightman et al., 1992), pause duration (Cooper & Paccia-Cooper, 1980; Scott, 1982) and intonation (Beckman & Pierrehumbert, 1986). Preboundary lengthening and pause duration were measured only with respect to the subject/VP boundaries indicated by hash marks above in the non-unit sequences and the corresponding place in the NP sequences, since these were the only two relevant phrasal boundaries in which the phonological context was identical between NP and NU versions of the two word sequences.

We examined preboundary lengthening by measuring the lengths of the vowels that occurred before a syntactic boundary in the NU sequences, and comparing them to the lengths of the same vowels in the corresponding NP sequences. Specifically, we examined the /u/ of “**gnu**/**new**”, and the /l̩/ of “peo**ple**”. As shown in Table 1, these acoustic measurements indicate that the length of the vowel before the subject/VP boundary in the NU sequence was longer than that of the same vowel in the syntactic-phrase-internal position in the NP sequence. This difference corresponds with the general finding that the vowels of syllables just prior to a syntactic boundary, particularly stressed vowels, are on average longer than the same vowel in other positions in the sentence (Beckman and Edwards, 1990; Cooper & Paccia-Cooper, 1980; Klatt, 1975; Wightman et al., 1992). However, there is a much greater difference in vowel length for the “gnu/new” than for the “people” set of stimuli, probably because “ple” is an unstressed syllable, and “gnu/new” is monosyllabic and hence necessarily stressed.

Insert Table 1 about here

We also measured the duration of the pause between syntactic phrases, and compared it to that within the phrase. As with preboundary lengthening, we compared only the “gnu/new” pair for the one sequence, and the “people” pair for the other sequence. As shown in Table 1, we found a measurable pause only at one location – namely the subject/VP boundary following “people” in passage 1 – and this pause was quite short.[4]

To describe the intonational contour of the utterances, we measured the highest and lowest points of f_0 (pitch) in each target sequence, and the highest and lowest points of f_0 in the “gnu/new” and “ple” syllables. As can be seen in Table 1, for both NP sequences the highest pitch point occurs during syllables in the first word of the phrasal sequence (on “new” and “ple”). These same syllables correspond to the preboundary syllables of the NU sequences: In these versions the highest pitch points occurs after the phrasal boundaries, on the first syllable of *watches*, “wa”, and on the vowel in *whole*. This may indicate a drop in pitch at the preboundary syllable in the NU sequences that is not found in the NP versions. However, the overall pitch excursion (i.e., the difference between the high and low points) in the preboundary syllables in the NU sequences is actually much smaller than in the same syllables in the NP versions. Referring to Figures

1-4, we see qualitatively that the NU sequence "gnu watches for men" contains a steeper rise in pitch to a high target on *watches* immediately following the subject/VP boundary than its NP counterpart. A similar high target on the verb is not detectable in the "people buy the whole" sequence.

Insert Figures 1-4 about here

In summary, the primary differences between our NP and NU stimuli were found in intonation and preboundary lengthening, not pause duration. With respect to intonation, we found two potential cues to the phrase boundary in the "gnu/new watches for men" stimuli, but only one of these cues in the "people by/buy the hole/whole" stimuli. Across the three cues, the differences we detected were more dramatic for the "gnu/new watches for men" stimuli than the "people by/buy the hole/whole" stimuli.[5] The latter result was unintentional in our construction and recording of the stimuli. The unintentional difference in the stimuli sets will be revisited in the discussion of the results of this experiment.

These measurements afford some interesting comparisons with the prosodic cues coincident with the clause boundaries in Nazzi et al.'s (2000) analogous set of stimuli. With respect to preboundary lengthening, the characteristics of our stimuli were similar to those of Nazzi et al. Both had longer vowels preceding the boundary in the NU sequences than the syntactically well-formed (phrase or clause) versions. With regard to intonation, there were again similarities between the potential cues to boundaries in the two stimulus sets. Nazzi et al. found a large rise in pitch at the boundary in the NU sequences that was not matched at the equivalent location in the clause sequences. A similar post-boundary rise in pitch was found in our stimuli. Also in both stimulus sets, a high pitch excursion was found on the first syllable of the well-formed sequence. It is worth noting that the rises in pitch on the post-boundary syllables in the Nazzi et al. NU sequences correspond to the beginnings of phrases – a verb phrase and an object noun phrase. In the first case, this corresponds exactly to the boundary of interest in our experiment.

The primary difference between our stimulus pairs and those from the Nazzi et al. experiments is revealed by the measurements of pause duration. Nazzi et al. found long pauses at the clause boundaries of the NU sequences (456 and 140 ms), which were an order of magnitude greater than the pauses at the same locations in their well-formed clause sequences (27 and 36 ms). By contrast, we found only a very small pause (18 ms) in just one of our NU sequences. Hence, if infants' segmentation of fluent speech is uniquely based on the distribution of long pauses, we should not expect to find results with our phrase stimuli that parallel Nazzi et al.'s results with clauses.

Design. The overall design of the experiment involved familiarizing each infant with the NP and NU versions of the same phonological word sequence, and then examining their preferences for listening to the passage containing the NP target sequence (the NP passage) compared with the passage containing the NU target sequence (the NU passage). To mitigate the influence of listening preferences irrelevant to the manipulation of interest, infants were divided into two groups that heard different phonologically matched target sequences during familiarization. Group 1 was familiarized with "**new watches for men"** (NP) and **"gnu watches for men**" (NU), whereas Group 2 was familiarized with "**people by the hole"** (NP) and **"people buy the whole**" (NU). Passage I contained the NP target sequence for Group 1 and the NU target sequence for Group 2, while the converse was true of passage II. For Group 1, a preference for passage I over passage II would indicate a processing advantage for the NP version. For Group 2, a preference for passage II over passage I would indicate the same advantage. Therefore, any overall preference for the NP passage over the NU passage across groups cannot be attributable to an overall preference for the passage itself (passage I or passage II), but must be attributed to an effect of the prosodic well-formedness of the familiarized sequences within the passages.

Apparatus. The stimuli were digitized at a 20 kHz sampling rate and stored on a MacIntosh G4 computer. The audio signal was then fed through a Yamaha audio amplifier to Cambridge Soundworks Ensemble II speakers.

The testing booth consisted of a 3-walled enclosure made of white pegboard panels, approximately 4.5 feet high, with white curtains that descended from the ceiling to meet the pegboard. The pegboard was backed by thick white cardboard to cover the holes, except for one large and two smaller openings in the front panel. The larger opening allowed a camera to record the session. A smaller opening allowed the experimenter to view the infant's headturns. Finally, a third opening (generally covered by the white curtain) allowed a secondary observer, such as a second parent or grandparent, to view the procedure. A chair was placed in the center of the booth, facing the front panel.

A light was attached at the center of each panel, at the approximate eye level of an infant seated on a caregiver's lap in the chair. The light on the front panel was green, while the lights on the side panels were both red. Each of the two loudspeakers was situated behind the two side panels, located directly behind the red light. The computer, response box, and other equipment were located behind the front panel, out of sight of the infant.

Procedure. A modified version of the Headturn Preference Procedure (Jusczyk, 1998b; Kemler Nelson et al., 1995) was used. Each infant was seated on the caregiver's lap on the chair in the middle of the testing booth. The experimenter was situated behind the testing booth, and observed the infant through the viewing hole. During the experiment, the orientation of the infant's gaze was recorded on the computer by means of a button box. All choices regarding the lights and auditory stimuli were made via computer program. Both the experimenter and the caregiver wore tight-fitting headphones that played continuous music to mask the auditory stimuli the infant heard. The overhead light was dimmed to make the panel lights more salient.

Each trial began with the front green light flashing to attract the infant's attention forward. When the infant looked forward, this light would extinguish, and one of the two side red lights (chosen randomly by the computer program used to run the experiment) would begin to flash. When the infant oriented to the side light, the auditory stimulus would play. This continued until either the infant looked more than approximately 30º away from the light for 2 consecutive seconds, or the entire stimulus file was complete. At this point, the side light would extinguish, the sound would stop, and the front green light would begin to flash in preparation for the next trial. The computer recorded the amount of time the infant was looking toward the light while the stimulus was playing. If the infant looked away for less than 2 s and then looked back again, the trial continued, but the amount of time spent looking away was not counted in the overall tally. This general procedure was the same for both familiarization and test trials.

The infants in each group were familiarized with their respective target sequences until they had accumulated at least 20 s total looking time to each target type (NP or NU), for a total of 40 s to the word sequence. Since each sound file was 16.0 s, this required a minimum of two trials for each sequence type. The infants were then immediately presented with the two different test passages, presented in four blocks, for a total of 8 test trials. Since the test trials consisted of 3 repetitions of the test passages, infants heard the target sequence between 0 and 3 times per test trial, depending on their orientation time during that trial. The order of presentation of the two passages was randomized within each block. The dependent measure was the average looking time across trials to each stimulus type.

Results

Mean looking times were calculated for each passage type for each familiarization group, and are shown in Figure 5. Twenty-three of the 32 infants had longer looking times to the passage containing the noun phrase target sequence (NP) than the passage with the non-syntactic-unit (NU). Infants listened to the NP passage an average of 10.5 s (SD = 4.0 s), and to the NU passage an average of 8.8 s (SD = 3.8 s). An ANOVA with two factors, one between-subjects (Familiarization Group: 1 or 2) and one within-subjects (Passage Type: NP or NU), showed a significant main effect of Passage Type (F (1, 30) = 5.81, $p < .05$) and a significant interaction between Familiarization Group and Passage

Type (F (1, 30) = 5.34, p < .05), but no main effect of Familiarization Group (F (1, 30) = 2.38, p > .10). A further analysis revealed a highly significant preference for the NP passage in Familiarization Group 1 (t (15) = 17.29, p < .001). These infants listened to the NP passage (containing “new watches for men”) an average of 10.5 s (*SD* = 4.5 s) and to the NU passage (containing “gnu watches for men”) an average of 7.1 s (*SD* = 2.9 s), with 15 out of 16 infants preferring the NP passage. However, there was no preference for either passage in Familiarization Group 2 (t (15) < 1). Infants in this group listened to both the NP passage (containing “people by the hole”) and the NU passage (containing “people buy the whole”) an average of 10.5 s (NP: *SD* = 3.6 s; NU: *SD* = 3.8 s), with only 8 of 16 infants preferring the NP passage. Hence, the overall effect of Passage Type was solely due to Familiarization Group 1.

To address the concern that infants might simply be more familiar with the prosodically well-formed NP target sequence than the NU target sequence within the passage, looking times were also inspected for the familiarization trials. Infants listened to the NP sequence an average of 10.4 s (*SD* = 3.3 s), and to the NU sequence an average of 11.2 s (*SD* = 3.0 s) per familiarization trial. Infants listened an average of 26.7 s (*SD* = 4.1 s) to the NP sequence and 27.0 s (*SD* = 4.2 s) to the NU sequence during the entire familiarization phase. An ANOVA with one between-subjects factor (Familiarization Group: 1 or 2) and one within-subjects factor (Sequence Type: NP or NU), revealed no significant effects (p > .1). A second analysis of these data was performed by inspecting the looking times on the first and last trials of the familiarization period. If either sequence type had a greater number of trials (this occurred in some cases when one stimulus had reached criterion but the other stimulus in the pair had not), the last block in which both stimulus types were played was used. These data were analyzed using an ANOVA with three factors, one between (Familiarization Group) and two within (Trial: first or last, and Sequence Type). There was a marginal main effect of Trial (F (1, 30) = 3.33, p = .08), due to greater looking times on the first trial, but no other effects approached significance (p > .1). Thus, infants did not differentially prefer or habituate to the sequence types during familiarization. Overall, then, the subsequent preference for the test passage containing the well-formed syntactic unit is not attributable to the sequence's greater acoustic familiarity as a result of experience in the familiarization phase of the experiment.

Discussion

Infants preferred to listen to the test passage containing the familiarized target sequence as a well-formed syntactic unit, a subject noun phrase. Since both test passages were themselves well-formed, prosodically and syntactically, it can be inferred that this preference is due to a prosodic effect on the infants' ability to recognize an embedded target sequence within the test passages. That is, the prosodic cues help the infants to better "find" the familiarized sequence in a passage of continuous speech when it is a prosodically well-formed noun phrase than when it is an ill-formed sequence. This effect is likely due to the fact that infants parse the passage into well-formed phrasal units when they process it, and thus are more likely to detect the recurrence of the well-formed unit than the ill-formed unit heard in familiarization.

An alternative account to consider is that the effect derives more indirectly –namely, as a consequence of the well-formed sequence being better remembered than the ill-formed sequence, and therefore better recognized when it recurs in the continuous speech passage. This second possibility – that infants could better remember the well-formed unit (NP) than the non-unit (NU) – is itself a potentially interesting finding since such an effect has never been demonstrated at the level of the phrase. All previous evidence that well-formed prosody confers a processing advantage has related to clausal segments (Mandel et al., 1994, 1996) rather than phrasal segments. However, we suggest that this is not an adequate account of the current findings. Given the nature of the experimental design, better memory for the words that constitute the well-formed sequence would confer no advantage on the passage that contained it, since the other passage contained exactly the same words in exactly the same order. Furthermore, since there was no difference in listening times or listening patterns (that is, preferences or rates of habituation) to the target sequences during familiarization, infants were equally familiar with both acoustic versions. Although the explanation of these results as a pure memory effect cannot be ruled out categorically, the more compelling conclusion is that infants do use prosodic cues marking phrases to segment passages of continuous speech into sub-clausal units.

Note, however, that the advantage for the passage containing the NP target sequence over the passage containing the NU target sequence occurred in only one of the two groups (Group 1) – namely, the group that heard "new watches for men" as the NP and "gnu watches for men" as the NU during the familiarization phase of the experiment. Infants in Group 2, familiarized with "people by the hole" as the NP and "people buy the whole" as the NU, did not preferentially attend to the passage with the familiar NP. Still, it is important to point out that this group of infants did not preferentially listen to the passage containing the familiar NU either (that is, passage I, which was the passage containing the familiar NP for Group 1). Group 2's mean listening times were identical for the two passages. Hence, the positive result in Group 1, in the context of the lack of preference in Group 2, cannot be explained by positing a mere preference for one passage over the other, independent of the sequences heard during familiarization. Rather, it constitutes the first evidence for an effect of phrase-level prosody on speech processing in infants.

This limitation on the generalizability of the finding, however, does raise an important question concerning the source of the difference between the two sets of stimuli. One interesting possibility is that the prosodic cues associated with the phrase boundary between “gnu” and “watches” were more salient than those between “people” and “buy”, such that the NU sequence was more clearly marked as crossing a phrase boundary in the comparison involving “gnu watches for men” (Group 1) than the comparison involving “people buy the whole” (Group 2). This suggestion is reminiscent of Gerken et al.'s (1994) conclusion that infants are sensitive to the subject/VP boundary only in contexts in which it is saliently marked with prosodic cues. Our acoustic analyses yielded results consistent with this account. Measurements of both preboundary lengthening and changes in intonation suggested stronger cues to the phrase boundary in the NU “gnu watches for men” than the NU “people buy the whole”. Gerken et al. also found that infants were sensitive to combinations of preboundary lengthening and changes in intonation, but not to one of these cues in isolation. Notably, the lack of a pause duration difference between NU and NP versions of the Group 1 stimuli did not negate infants' preference for the NP target. This cue may well play an important role at the level of clauses, as in Nazzi et al. (2000), where the cue is highly salient, but it is apparently not necessary at the phrase level. Indeed, the only cue that favored the stimuli of Group 2 was the small and non-salient pause after “people”, which was evidently ineffective.

A remaining possibility is that, in addition to the prosodic effect, the infants had an overall preference for passage I. In the case of Group 1, the two effects, the basic preference for passage I, as well as the better recognition of the well-formed target sequence, would combine for an overall preference for passage I. However, in the case of Group 2, these effects would be in conflict, and no overall preference for either passage would be evident. To investigate this further, we ran a control experiment in which 16 infants were exposed to two versions of an irrelevant sequence (the second pair of familiarization sequences to be used in Experiment 3) during familiarization, and then tested on the passages from Experiment 1. We found no suggestion of an overall preference for either passage (*t* (15) < 1). Mean listening time to passage I was 10.8 s (*SD* = 4.1 s) and to passage II, 9.8 s (*SD* = 3.4 s). Nine of 16 infants preferred passage I, and 7 preferred passage II. An ANOVA with two factors, one between-subjects (Familiarization Group: Group 1 or control) and one within-subjects (Passage: I or II), revealed a trend toward an interaction between Familiarization Group and Passage (*F* (1, 30) = 2.88, *p* = .10), with a stronger preference in Group 1. Moreover, an inspection of the proportions of infants preferring passage I over passage II indicated that the proportion was larger in Group 1 than in the control group (_^{2} (1) = 4.16, *p* < .05). There were no hints of a difference in looking patterns between Group 2 and the control group when parallel tests were applied.

Accordingly, the lack of a reliable preference in the control group, in combination with evidence that the preference was stronger in Group 1 than the control group, suggests that a sheer preference for one of the test passages over the other did not contribute to the critical tendency shown by Group 1 to attend longer to the passage containing the familiarized NP sequence than the one containing the familiarized NU sequence. In any case, the fact that Group 2 showed no hint of a preference for passage I, which constituted their NU passage, over passage II makes clear that the reliable preference for the NP passage exhibited by Group 1 cannot be solely the consequence of a preference for one passage over the other, independent of the difference in well-formedness of the sequences the infants heard during familiarization.

Overall, the principal findings in Experiment 1 suggest that 9-month-olds do sometimes use phrase-level prosody in parsing continuous speech. This is the age at which Jusczyk et al. (1992), using the pause-insertion paradigm, first found evidence that infants were sensitive to the prosodic cues associated with phrase boundaries. By contrast, the 6-month-olds tested by Jusczyk et al. failed to show the sensitivity. Yet, at the level of the clause, 6-month-old infants have been shown to use prosody in speech processing (Nazzi et al., 2000), and even 2-month-old infants have been shown to be sensitive to the prosodic cues associated with clause boundaries (Mandel et al., 1994, 1996).

An earlier sensitivity to clause-level prosody than phrase-level prosody is understandable, because phrase-level prosody is more subtly marked and more variable than clause-level prosody. Indeed, phrase-level prosody varies dramatically across languages. For example, Korean phrasal prosody is very different from English phrasal prosody (Beckman & Jun, 1996), primarily because Korean is an accentual language and English is an intonational one. Infants might therefore need extra time to “tune” their acquisition system to the prosodic properties relevant in the ambient linguistic environment. In addition, the variability of prosodic properties is also different across languages. For example, English prosody is more variable than Dutch prosody (e.g., Neeleman & Reinhart, 1998). So English-learning infants might need more time to adjust to English prosodic properties than Dutch children to Dutch prosody. The progression to language-specific sensitivities is also found in other domains of speech perception at about this time, toward the end of the first year (for example, phoneme contrasts, as shown by Werker & Tees, 1984).

Accordingly, we might expect that 6-month-olds would not show the same advantage for well-formed phrasal targets over non-units in the recognition task of Experiment 1. However, given that infants as young as 7.5 months can segment speech into smaller, more language-specific units than the clause – namely, words (Juscyzk & Aslin, 1995) – and can also segment into clauses as young as 6 months (Nazzi et al., 2000), it is plausible to suppose that 6-month-olds, like 9-months-olds, might use phrase-level prosodic cues in parsing the speech stream. In Experiment 2, we examine whether these younger infants demonstrate the same sensitivity to phrase-level prosody as 9-month-olds. A positive result would constitute the first indication that infants are sensitive to phrase-level prosody by 6 months of age. It would further imply that 6-month-olds use this prosodic information in processing speech. This information is important for understanding the relationship and timing of sensitivity to phrase-level prosody in relation to the development of other aspects of language, such as lexical knowledge. It also has implications regarding the extent to which this sensitivity depends on the acquisition of language-specific knowledge of prosodic correlates of syntax.

Experiment 2

In Experiment 1, 9-month-olds demonstrated a preference for the passage with the well-formed version of a noun-phrase (NP) word sequence over a phonologically matching non-unit (NU) version after familiarization with both. This suggests that 9-month-old infants sometimes use phrase-level prosodic cues in segmenting continuous speech. If 6-month-olds also use prosodic cues in this way, then they too should show a preference for the passage containing a familiarized NP target sequence over a passage containing a phonologically identical familiarized NU target sequence.

Method

Participants. Thirty-two infants (19 males, 13 females) from monolingual English-speaking households in the Baltimore area were tested. The infants had an average age of 26 weeks and 5 days (range = 25 weeks; 4 days to 28 weeks; 4 days) or approximately 6 months. Nine additional infants were tested but not included in the study. Seven infants were not included for the following reasons: crying or fussing (4), average listening time less than 3 s (1), experimenter error (1), one parent spoke British English (1). Another 2 infants were excluded because they did not meet a listening criterion of more than 7 s on two separate trials for each stimulus type during the test phase.

Stimuli, Design, Apparatus, Procedure. The stimuli, design, apparatus and procedure were identical to those in Experiment 1.

Results

Mean looking times were calculated for each passage type for each group, and are shown in Figure 6. Twenty-two of the 32 infants had longer looking times to the passage containing the noun phrase (NP) target sequence than the passage with the syntactic non-unit (NU). Infants listened to the NP passage an average of 13.3 s (*SD* = 5.2 s), and to the NU passage an average of 11.4 s (*SD* = 4.9 s). An ANOVA with one between-subjects factor (Familiarization Group: 1 or 2) and one within-subjects factor (Passage Type: NP or NU), revealed a significant main effect of Passage Type (F (1, 30) = 6.47, $p < .05$) and a significant interaction between Familiarization Group and Passage Type (F (1, 30) = 9.76, $p < .005$), but no main effect of Familiarization Group (F (1, 30) < 1. A further analysis showed a highly significant preference for the passage containing the NP target sequence (the NP passage) for Group 1 (t (15) = 11.78, $p < .005$), with infants listening to the NP passage an average of 13.7 s (*SD* = 5.1 s) and the NU passage an average of 9.5 s (*SD* = 4.4 s). Fifteen of 16 infants in Group 1 preferred the NP passage. However, there was no preference for either passage in Group 2 (t (15) < 1), with infants listening to the NP passage an average of 12.9 s (*SD* = 5.5 s) and the NU passage an average of 13.3 s (*SD* = 4.7 s). Seven of 16 infants in Group 2 preferred the NP passage. Hence, as with the 9-month-olds, the overall effect of Passage Type was solely due to Group 1, the group familiarized with “new watches for men” as the NP target sequence and “gnu watches for men” as the NU target sequence.

Again, to address the possibility that infants might be more familiar with the prosodically well-formed NP sequence than the NU sequence, looking times were calculated for the familiarization trials. Infants listened to the NP target sequence an average of 12.4 s (*SD* = 2.8 s), and to the NU target sequence an average of 12.3 s (*SD* = 3.3 s) per trial. Infants listened an average of 27.3 s (*SD* = 3.7 s) to the NP target sequence and 27.2 s (*SD* = 4.0 s) to the NU target sequence during the entire familiarization phase. An ANOVA with one between-subjects factor (Familiarization Group: 1 or 2) and one within-subjects factor (Sequence Type: NP or NU), revealed a significant effect of Familiarization Group (F (1, 30) = 8.031, $p < .01$), with Group 2 listening significantly longer (58.6 s, *SD* = 4.0 s) overall to the familiarization stimuli than Group 1 (52.5 s, *SD* = 6.9 s). There was no significant main effect of Sequence Type or interaction between Familiarization Group and Sequence Type ($p > .5$). As with Experiment 1, a second analysis of these data was performed using an ANOVA with one between-subjects factor (Familiarization Group) and two within-subjects factors (Trial: first or last, and Sequence Type). There was again a significant main effect of Familiarization Group (F (1, 30) = 8.347, $p < .01$), and also, as with the 9-month-olds, a marginal main effect of Trial (F (1, 30) = 3.144, $p = .09$) due to greater looking time on the first trial than the last. Importantly, no other effects approached significance ($p > .1$). Thus, like the 9-month-olds, the 6-month-old infants did not differentially prefer or habituate to the sequence types. Again, the preference for the test passage containing the well-formed syntactic unit is not attributable to the sequence's greater attractiveness or acoustic familiarity in isolation.

Discussion

The 6-month-olds, like the 9-month-olds in the first experiment, showed a preference for the passage containing the familiarized noun phrase (NP) over the passage containing the NU sequence. Also like the 9-month-olds, the effect was isolated to the group that heard the “new/gnu watches for men” sequences during familiarization. Since the 6-month-olds listened about equally to the two types of sequences during familiarization, it is not plausibly a greater familiarity with the NP version of the sequence that produces the advantage for the test passage that contains it. Accordingly, the findings of Experiment 2 constitute the first evidence that 6-month-olds are sometimes sensitive to prosodic cues to syntactic phrases, and furthermore, the first evidence that they sometimes use these cues in speech processing.

The current finding suggests that some sensitivity to cues that *may* be language specific (specifically, prosodic characteristics of syntactic phrases in the speech stream) exists by the age of 6 months, at least for infants learning American English. This early use of phrase-level cues may be due in part to the fact that some cues to phrase boundaries in English may be similar to cues to clause boundaries, a less language-specific prosodic unit. However, without a more exact characterization of which cues infants are using, this is difficult to determine. Whether the sensitivity reflects language-general characteristics typical of the clause as well as the phrase, or language- and phrase-specific characteristics to which the infant has already become attuned, is an open question. Certainly, there is evidence of other language-specific sensitivities developing prior to 6 months. Indeed, infants as young as a few days old are capable of discriminating between their native language and other languages (Mehler et al., 1988), and discriminating two rhythmically distinct languages, even in the absence of all but rhythmic cues (Ramus, 2002).

In sum, the preference pattern of the 6-month-olds was similar to that of the 9-month-olds: An effect was found in each case in the group that heard the “new/gnu watches for men" familiarization sequences, with no suggestion of a reliable preference in the other group, familiarized with the "people by/buy the hole/whole" sequences. These data add further support to the idea that the stronger set of prosodic cues differentially associated with the stimuli in Group 1 made the difference between the noun phrase and non-unit targets more salient in Group 1 than it was for Group 2.

The contrast between the performances of Group 1 and Group 2 also suggests a possible explanation for the discrepancy between our findings and those of Jusczyk et al. (1992), who did not find evidence that 6-month-olds are sensitive to the subject/VP boundary. It may be that the difference is due to a difference between the two experimental tasks: 6-month-olds might be sensitive to prosodic cues marking phrases in the context of processing normal speech, but not in the context of detecting the insertion of artificial pauses. However, the discrepancy in sensitivity within our own stimuli suggests a different reason. It may be that the cues present in the Jusczyk et al. stimuli were simply weaker, or not of the right type to generate a significant preference in the younger infants.[6] Whether infants, either at 6 months or at 9 months, use phrase-level prosodic cues in processing speech may depend on whether there is a sufficient confluence of evidence.

The next experiment examines whether there are other possible constraints on the ability of 6-month-old infants to use speech cues to phrasal units in processing. To this point, our evidence has involved the specific contrast between non-unit sequences and well-formed *noun phrases*. As a crucial step in understanding exactly *where* prosodic cues are used, we ask whether other lexically headed phrases – in this case, verb phrases (VPs) – can also be successfully parsed in continuous speech.

Experiment 3

The first two experiments produced evidence that infants sometimes use prosodic cues at a phrasal level in their representation of speech. But the evidence also suggests that the infants are not always able to detect or use prosody as a means of segmenting clauses into phrases. In addition, our acoustic analyses of the stimuli revealed that the strength of the cues associated with the phrase boundaries was variable, in accord with previous claims. Overall, infants appear to be sensitive to prosodic cues to phrasal units at some ages but not others (Jusczyk et al., 1992), in some phonological/syntactic contexts but not others (Gerken et al., 1994), and in some acoustic contexts but not others (Gerken et al., 1994; the current Experiments 1 and 2). In order to develop a fuller picture of what phrase-level prosody might afford the bootstrapping language learner, it is necessary to have a clearer understanding of the contexts in which phrase-level prosody is, and is not, used.

The third experiment begins to address this issue by looking at another phrasal constituent within the clause – namely, the verb phrase (VP). This constituent is an important syntactic unit, on a par with the subject noun phrase. However, even though NPs and VPs are both lexically headed phrases, we might expect VPs to be prosodically distinct from subject NPs because (a) they usually do not occur clause-initially, and (b) they often contain an object NP that bears a nuclear pitch accent. In fact, in the default intonation pattern of English transitive sentences that do not have contrastively focused elements, one generally finds nuclear stress on the most deeply embedded constituent of the sentence, usually the direct object (Cinque, 1993; see also Beckman & Pierrehumbert, 1986). This is also the canonical location of information new to the discourse, and hence often receives a high pitch accent, making it both prosodically and informationally salient (Pierrehumbert and Hirschberg, 1990; Nooteboom & Terken, 1982). This stress pattern might act as a cue to verb phrases, not available in the case of noun phrases. Looking at the verb phrase allows us to begin addressing the generalizability of the infants' sensitivity to various phrasal syntactic units in different prosodic contexts.

Method

Participants. Thirty-two infants (12 males, 20 females) from monolingual English-speaking households in the Baltimore area were tested. The infants had an average age of 26 weeks and 0 days (range = 24 weeks; 6 days to 28 weeks; 0 days) or approximately 6 months. Thirteen additional infants were tested but not included in the study. Ten infants were not included for the following reasons: crying or fussing (6), average listening time less than 3 s (1), failure to look to the lights (1), infant spit up (1), parents spoke Scottish English (1). Another 3 infants were excluded because they did not meet a listening criterion of more than 7 s on two separate trials for each stimulus type during the test phase.

Stimuli. The stimuli were produced by the same female native speaker of American English who produced the stimuli for Experiments 1and 2. They were produced under similar conditions and in the same manner as they were for the stimuli in Experiment 1. The following passages were recorded:

(III) Inventive people **design telephones** at home. A fresh idea with **promise # surprises** no-one who works there.

(IV) The director of **design # telephones** her boss. New developments **promise surprises** for their old buyers.

As with the stimuli in Experiments 1 and 2, each passage contained two target sequences, one well-formed syntactic unit (in this case, a simple verb phrase (VP) including a verb and its direct object) and one syntactic non-unit (NU), again crossing the subject/VP boundary. Like the stimuli in Experiment 1, both passages contained the same two phonological word sequences, but the syntactic context was reversed between the two. In Passage III, "**design telephones**" constituted a well-formed VP, while "**promise # surprises**" was a NU. By contrast, in Passage IV, "**promise surprises**" constituted a well-formed VP, while "**design # telephones**" was a NU.

The relevant syntactic boundaries in these pairs of stimuli were similar, but differed in some important ways, from the noun phrase stimuli in Experiments 1 and 2. In the noun phrase stimuli, the boundaries of the well-formed noun phrase occurred between an adjunct separated by a comma in the text passage, and the beginning of the verb phrase (the subject/VP boundary). The non-unit crossed the subject/VP boundary. In the verb phrase stimuli, the boundaries of the well-formed unit were between the subject/VP boundary and an adjunct to the verb phrase, an arguably subtler ending. However, the non-unit, as with the noun phrase stimuli, spanned a subject/VP boundary. In addition, these familiarization stimuli contained only open class items, whereas the stimuli in Experiment 1 contained both open and closed class items. This may be important because it will affect word stress.

The familiarization stimuli consisted of the four target sequences, extracted from the passages. Each sequence was 1.1-1.3 seconds long. Silence was added to the end of each sequence up to 2.0 s, and then the sequence was repeated 8 times, to make a final sound file of 16.0 s.

The test stimuli again consisted of two recorded passages. The durations of these passages were 7.3 s and 7.4 s respectively. Silence was added to the end of each of these passages up to 8.0 s, and then each passage was repeated 3 times, to make a final sound file of 24.0 s for each test passage.

These stimuli were analyzed in the same ways as the stimuli in Experiment 1. Intonation was measured for the entire target sequence, while preboundary lengthening and pause duration were measured only with respect to the subject/VP boundaries, indicated by hash marks above in the NU sequences, and the corresponding places in the VP sequences.

Acoustic measurements of pause duration as well as the word-final vowels (/I/ of *promise* and /aɪ/ of *design*) are shown in Table 2. Both preboundary vowels are longer than the corresponding phrase-internal vowels, although this increase is greater in the case of the "design telephones" stimuli set. In addition, in both cases the pause duration is longer at the subject/VP boundary than it is phrase internally, with a slightly greater increase in the case of "design telephones". Still, despite the existence of pauses at the subject/VP boundaries of the NU sequences, they are extremely small. Thus, with regard to pause duration, the verb phrase stimuli are much like the noun phrase stimuli used in Experiments 1 and 2, and unlike the stimuli in the Nazzi et al. (2000) study, in which there were large pauses at the relevant syntactic (clause) boundary.

As with the stimuli in the previous two experiments, we measured the highest and lowest pitch points of each target sequence, and the highest and lowest points of the preboundary syllables and their counterparts in the VP versions. As shown in Table 2, for both NU sequences the NU preboundary syllables ("sign" and "mise") have *lower* apex pitch peaks than their VP counterparts, and the preboundary pitch excursions are greater in the NU versions of the sequences than the VP versions (unlike what was found with the NP stimuli). Referring to Figures 7-10, we see that the NU sequence "design telephones" contains a steep rise in pitch immediately post-boundary (similar to the Nazzi et al. stimuli, as well as the stimuli of Group 1 in the previous experiments), whereas its VP counterpart shows a more gradual rise. However the NU does not show such a steep rise (if anything, there is a decline) post-boundary, while the VP version has a very large rise on the final syllable of "surprises". The intonational properties associated with these verb phrase stimuli are clearly very different from those of the noun phrase stimuli in Experiments 1 and 2.

Insert Figures 7-10 about here

Yet, like the noun phrase stimuli, our measurements found that a more salient prosodic boundary in the NU sequence followed the stressed syllable, while a less salient boundary followed the unstressed syllable.[7] One therefore might expect a larger effect in the group familiarized with "design telephones" rather than "promise surprises". This would be in keeping with the findings of Experiments 1 and 2, in which only the group exposed to the more salient cues showed a preference for the well-formed phrasal unit.

However, both verb phrases in our stimuli conformed to the canonical intonation pattern for verb phrases in English, which may serve as a cue for both groups. In addition, although the break in the NU sequence was more salient for "design telephones", there was still the critical difference in preboundary lengthening and pause duration for the two versions of "promise surprises". Therefore, both groups might show a reliable preference.

Insert Table 2 about here

Design. As in Experiments 1 and 2, there were two groups of infants familiarized with different pairs of phonologically matching sequences. Familiarization Group 1 heard the "**design telephones**" stimuli (both the VP and NU versions) during familiarization, while Familiarization Group 2 heard the two versions of "**promise surprises**". All infants then heard both test passages. For Familiarization Group 1, the target word sequences heard during familiarization were a well-formed VP in Passage III and a NU in Passage IV. For Familiarization Group 2, the target word sequences heard during familiarization were a well-formed VP in Passage IV and a NU in Passage III.

Apparatus, Procedure. Both the apparatus and procedure were identical to those used in Experiments 1 and 2.

Results

Mean looking times were calculated for each passage type for each group, and are shown in Figure 11. Twenty-one of the 32 infants had longer looking times to the passage containing the VP sequence with which they were familiarized, than to the passage with the NU sequence with which they were familiarized. Infants listened to the VP passage an average of 14.1 s (*SD* = 5.0 s), and to the NU passage an average of 11.9 s (*SD* = 3.3 s). An ANOVA with one between-subjects factor (Familiarization Group: 1 or 2) and one within-subjects factor (Passage Type: VP or NU), showed a significant main effect of Passage Type (F (1, 30) = 4.49, $p < .05$) but no significant interaction between Familiarization Group and Passage Type (F (1, 30) < 1), and no main effect of Familiarization Group (F (1, 30) < 1). The mean listening times for Group 1, familiarized with the "design telephones" sequences, were 14.1 s (*SD* = 5.3 s) to the VP passage and 12.0 s (*SD* = 3.1 s) to the NU passage. In Group 1, 10 out of 16 subjects listened longer to the passage containing the VP sequence. The mean listening times for Group 2, familiarized with the "promise surprises" sequences, were 14.0 s (*SD* = 4.9 s) to the VP passage and 11.8 s to the NU passage (*SD* = 3.6 s). In Group 2, 11 out of 16 subjects listened longer to the passage containing the VP sequence.

Looking times were also calculated for the familiarization trials, in parallel with Experiments 1 and 2, to check that infants were not more attracted to, or familiar with, the prosodically well-formed VP target sequence than the NU sequence. Infants listened to the VP target sequence an average of 12.6 s (*SD* = 2.7 s), and to the NU target sequence an average of 12.1 s (*SD* = 2.8 s) per trial. Infants listened an average of 27.2 s (*SD* = 3.2 s) to the VP target sequence and 28.6 s (*SD* = 4.7 s) to the NU target sequence during the entire familiarization phase. An ANOVA with one between-subjects factor (Familiarization Group: 1 or 2) and one within-subjects factor (Sequence Type: VP or NU) revealed no significant effects ($p > .1$). As with Experiments 1 and 2, an analysis of the pattern of looking times was performed using an ANOVA with one between-subjects factor (Familiarization Group) and two within-subjects factors (Trial: first or last, and Sequence Type). No effects approached significance ($p > .1$). Thus, as in the previous experiments, the infants did not differentially prefer or habituate to the sequence types during the familiarization phase.

Discussion

Infants showed a preference for the passage containing the target sequence as a well-formed verb phrase (VP) over the passage containing the syntactic non-unit (NU). In combination with the results of Experiment 2, this may suggest that 6-month-olds are independently attentive to prosodic cues that mark the verb phrase and the noun phrase as syntactic units, or it may suggest again that infants are attentive to the subject/VP boundary that interrupts the syntactic non-unit. As before, the infants did not significantly prefer or habituate to one type of sequence during familiarization, so the preference for the passage containing the VP over the passage containing the NU is not plausibly the result of the greater acoustic attractiveness or familiarity of the VP sequence in isolation.

In this third experiment, there was no evidence of an interaction involving the specific sequences with which the infants were familiarized. The preferences demonstrated by the infants were not reliably different for the two groups. Based on the acoustic analyses, one might have expected that Group 1 would show a greater preference, since there seemed to be a stronger prosodic difference between the verb phrase and non-unit sequences for this group. Instead, across groups, there was an overall preference for the passage with the well-formed phrasal target, suggesting that, if the subject/VP boundary within the non-unit sequence is the source of the effect, it was salient for both groups. In summary, while the acoustic/prosodic properties associated with the verb phrase stimuli were quite different from those of the noun phrase stimuli, infants showed evidence of using the cues in both cases in order to parse the speech stream into major phrasal units.

General Discussion

The results reported in this paper constitute the first evidence that syntactically-influenced prosodic cues may affect infants' recognition of major phrasal units embedded within continuous passages of speech. In several significant instances, infants preferred passages containing a familiar phonological sequence when this sequence constituted a well-formed phrasal unit (NP or VP) than when it crossed the subject/VP boundary and constituted a syntactic non-unit. Furthermore, while previous work suggested that infants might not be sensitive to phrase-level prosody until the age of 9 months, the current work demonstrates that infants as young as 6 months of age sometimes differentiate between well-formed noun or verb phrases and syntactic non-units, based on phrase-level prosodic characteristics.

These findings have important consequences for our understanding of the development of tools for language acquisition in the first year of life. Previous work has shown that infants are sensitive to prosodic characteristics of their language within the first few days of life (Mehler et al., 1988), and that by the age of 2 months, infants are using prosodic information in their organization of speech input in memory (Mandel et al., 1994, 1996). By 4.5 months, infants are sensitive to the prosodic correlates of clauses (Jusczyk, 1997), and by 6 months, they are capable of using prosodic information regarding clausal units in their processing of continuous infant-directed speech (Nazzi et al., 2000). At 7.5 months, infants begin to use prosodic information to segment individual words from the speech stream, and at 10.5 months, they are using other kinds of cues, such as phonotactic information, to segment words (Jusczyk et al., 1999).

The findings of Jusczyk et al. (1992) and Gerken et al. (1994) suggested that infants were not sensitive to prosodic correlates of phrasal units until 9 months of age. This is *after* infants show evidence of segmenting individual words from the speech stream (Jusczyk and Aslin, 1995; Jusczyk et al., 1999). The current work supports, instead, the idea that infants are sensitive to phrase-level prosody *before* they use prosodic information for word segmentation. This suggests a picture of the development of prosodic competence that proceeds from larger, more language-general units (the clause), through intermediate units (phrases) to smaller, more language-specific units (the word). Similar progressions from larger to smaller have been suggested in the area of face recognition (Ghim & Eimas, 1988) and scene processing more generally (Fricka, Colomboa, & Allen, 2000). From the infants' perspective, this progression might be a smooth one. It has not been shown that infants distinguish at the outset between clauses and phrases, or phrases and words. What to the adult listener is easily identified as a syntactic unit might, for the infant, initially simply be a cohesive unit for analysis.

The use of prosody in infants' processing of speech is generally viewed as enabling the parsing of the input into smaller syntactically relevant units, as suggested in Jusczyk (1998a). Yet, it has not been established exactly how this parsing process might aid in the acquisition of syntax and other aspects of the grammar. Certainly, it would be unreasonable to propose that infants simply read syntactic trees off the acoustic signal from the prosodic cues, nor is this what Jusczyk intended by his proposal. There is simply too much variability in the cues, and it is not clear to what extent (if at all) such cues correlate with more minor phrase boundaries than those examined here. Instead, it is likely that this parsing process, together with other sources of information present in the speech stream, such as the location of function morphemes, allows the infant to more easily detect the regularities of the grammar. For example, recent evidence indicates that, in languages with rich agreement systems, agreement morphology may be used by infants to parse phrasal domains from continuous speech (Blenn, Seidl, & Hoehle, in press). Jusczyk (1998a) puts it this way: "[parsing] provides some bounds for restricting the search for regularities in the way that certain patterns are distributed within the input" (p. 293). This might permit the infant to make important comparisons across sentences, or even phrases, as suggested by previous work (Gerken, 1996; Gerken et al., 1994; Jusczyk & Kemler Nelson, 1996). Other sources of information, such as those proposed by Fisher and Tokura (1996b), may also come in to play in this undertaking.

The current findings suggest that 6-month-olds sometimes have access to, and employ in fluent speech processing, the basic information necessary to deploy such a "divide and conquer" strategy, and that they have the means to divide the speech stream into syntactic units smaller than the clause. A strong interpretation of these results, by which stretches of speech corresponding to noun phrases and verb phrases (or their prosodic correlates) are psychological units for infants, would evidently have important consequences for infants' ability to pull out syntactically relevant units from the speech stream. Even a weaker interpretation of the findings, by which infants are simply sensitive to the subject/VP boundary that was present in all the non-unit stimuli – such that its presence disrupts infants' ability to recognize the recurrence of the sequence in the passage – would still confer upon infants the tools to parse the speech stream into phrase-level units, based on this phrase-boundary marking in conjunction with clause-boundary marking. The "divide and conquer" proposal (Jusczyk, 1998a) suggests that the ability of infants to parse the speech stream into phrasal units would greatly help infants to detect many different kinds of regularities in the input – including phonotactic regularities, the presence and locations of function morphemes and other lexical items, and word-level prosodic patterns.[8] Furthermore, these benefits could accrue even given an imperfect correspondence between syntactic phrases and the units delimited by prosodic cues.

There remain two important and interrelated issues with respect to infants' use of phrase-level prosodic cues in speech processing. The first pertains to the nature of the input. It is evidently not enough to say that phrase-level prosody is variable, or that phrase-level prosody is correlated with syntactic boundaries. To gain perspective on *how much* and *what* infants might gain by paying attention to these cues, it is necessary to have a more exact characterization of the strength and type of cues in both infant-directed and adult-directed speech, and their consistency with respect to various syntactic boundaries. This may also increase our understanding of whether infants distinguish between phrase boundaries and clause boundaries — a question that the current findings do not directly address. The second issue pertains to the nature of infants' sensitivity to these features of the input. While a controlled manipulation of the various cues is in order, it is clear that large pauses, such as those found between clauses but more rarely between phrases, are not the only acoustic features that infants are using to process speech. So far, the evidence suggests that infants are capable of detecting and using phrase-level prosody in some circumstances but not others, and that a confluence of cues may be necessary in order for infants to use these cues (Gerken et al., 1994; Jusczyk et al., 1992; the current study). It is possible that the need for multiple cues helps infants to avoid falsely identifying non-syntactic prosodic information (such as emphatic stress) as cues to syntactic boundaries. A detailed characterization of the types of prosodic cues to which infants are sensitive and the circumstances under which infants use these cues to process and encode speech would constitute major steps forward.

Although these important issues remain to be resolved, the current study is the first to suggest that 6-month-olds, as well as 9-month-olds, are sometimes sensitive to, and indeed sometimes use, phrase-level prosodic cues in processing continuous speech. While it is clear that there are limitations on this ability (and we have pointed to some of these), the demonstration that such an ability exists at all at such a young age has important consequences for our notions of the developmental course of infant speech processing, as well as other aspects of language acquisition. Infants' ability to parse the speech stream could provide a foundation for many facets of language acquisition – not only learning syntax, but also word segmentation and morphological acquisition.

Acknowledgements

Support for this research was provided by a Research Grant from NICHD (15795) and a Senior Scientist Award from NIMH (01490) to PWJ, a National Science Foundation Graduate Research Fellowship to MS, an IGERT postdoctoral fellowship to AS, and a Swarthmore College sabbatical award to DGKN. Thanks are due to Ann Marie Jusczyk and Natasha Henline for their help and support in recruiting and running subjects, to Stefanie Shattuck-Hufnagel for advice on the presentation of the intonation data and discussion of prosodic boundary cues, to Edward Kako for fruitful ideas about how phrasal segmentation might aid language acquisition, and to Eon-Suk Ko for help with the ToBI coding. We wish to thank four anonymous reviewers for their insightful comments and suggestions on a previous draft of this manuscript. Appreciation is also due to Barbara Landau, Karen Arnold, Elizabeth Johnson, George Hollich, and several others too numerous to list, whose support has facilitated the completion of this and other projects following the death of PWJ.

Footnotes

[1] Although it was necessary to choose an order of authorship for publication, this work has been a true collaboration, and each author's contribution to the project has been equally important.

[2] The question of whether prosodic cues to phrases are different from those to clauses is complicated by the indirect relationship between syntactic constituents (e.g., noun phrase, verb phrase) and prosodic constituents (e.g. Intonational Phrase, Phonological Phrase). The studies reported here compare across prosodic, rather than syntactic, levels of representations. Current theories (e.g. Selkirk, 1996; Truckenbrodt, 1999; Seidl, 2001) suggest that, at least to some degree, the rules that map prosodic domains to syntactic ones ignore functional projections and only take note of lexically-headed projections and their boundaries. Therefore, lexically-headed phrases such as the ones used in our study correspond to a greater degree with prosodic constituents than functionally-headed phrases (such as determiner phrases, complementizer phrases or tense phrases). Although there is not a perfect correlation, for the purpose of the current discussion it is reasonable to equate Phonological Phrases with lexically-headed syntactic phrases and Intonational Phrases with clauses in order to discuss what these studies have to say about the prosodic differences between clauses and phrases.

[3] Since the speaker initially mis-pronounced the word "gnu" with a /g/ sound (and was asked to re-record this passage), it is unlikely that she caught on to the exact pairing between the stimuli, although she may have noticed some similarity across the items.

[4] In fact, when ToBI analyses (a system of marking prosodic information in speech, Beckman & Hirschberg, 1993) were performed on these data by two independent

coders, both coders chose a break index of 1 for this boundary, suggesting that this pause was not perceived by the coders.

[5] ToBI analyses of these data confirmed this basic finding.

[6] A third possibility also presents itself. It is noteworthy that the stimuli used in the current study are all bare NPs rather than DPs (phrases containing a determiner like “the”, e.g. “new watches for men” rather than “the new watches for men”). In contrast, many of the subject NPs used in the Jusczyk et al. (1992) study were headed by determiners (DPs). These may be more difficult to detect due to the presence of an acoustically unstressable element (a determiner) occurring phrase-initially, hence making Jusczyk et al.’s task more difficult than ours.

[7]Again, ToBI analysis concurred with these acoustic measurements.

[8] A provocative possibility regarding how infants might use prosodic information in conjunction with the parsing of phrases to obtain information about syntax is that of the Rhythmic Activation Principle (Nespor, Guasti, & Christophe, 1996). This hypothesis suggests that infants might be able to discover the branching direction of the target language or the location of phrase heads by relying on prosodic prominence relations in the input language. This strategy would provide infants with crucial information for discovering the word order of their language, assuming that they are capable of locating phrase edges in the speech stream. The current study suggests that infants might have access to such phrasal information well before the end of the first year.

References

Beckman, M.E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston & M.E. Beckman (Eds.), Papers in laboratory phonology I: Between the grammar and the physics of speech (pp. 152-178). Cambridge: Cambridge University Press.

Beckman, M. E. & Edwards, J. (1992). Intonational categories and the articulatory control of duration. In Y. Tohkura, Eric Vatikiotis-Bateson, and Y. Sagisaka, (Eds.), Speech, perception, production and linguistic structure (pp. 359-376). Tokyo: OHM Publishing Co.

Beckman, M. E., & Hirschberg, J. (1993). The ToBI annotation conventions. Unpublished manscript, retrieved from http://ling.ohio-state.edu/Phonetics/etobi_homepage.html.

Beckman, M. E., & Jun, S.-A. (1996). K-ToBI (Korean ToBI) labeling convention, Version 2. Unpublished manuscript, Ohio State University and University of California: Los Angeles.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. Phonology Yearbook, 3, 255-309.

Blenn, L., Seidl, A., & Hoehle, B. (in press). Recognition of phrases in early language acquisition: The role of morphological markers. In Proceedings of the 26th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. Cognition, 61, 93-125.

Brown, R. (1973). A first language: The early stages. Cambridge, MA: Harvard University Press.

Cho, T., & Keating, P. (1999). Articulatory and acoustic studies of domain-initial strengthening in Korean. UCLA Working Papers in Phonetics, 97, 100-138.

Cinque, (1993). A null theory of phrase and compound stress. Linguistic Inquiry, 24, 239-298.

Cooper, W. & Paccia-Cooper, J. (1980). Syntax and speech. Cambridge, MA: Harvard University Press.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of vowel-initial syllables as a function of prosodic structure. Journal of Phonetics, 24, 423-444.

Fisher, C., & Tokura, H. (1996a). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. Child Development, 67, 3192-3218.

Fisher, C., & Tokura, H. (1996b). Prosody in speech to infants: Direct and indirect cues to syntactic structure. In J. Morgan & C. Demuth, (Eds.), Signal to Syntax (pp. 343-364). Mahwah, NJ: Erlbaum.

Fricka, J.E., Colomboa, J., & Allen, J.R. (2000). The temporal sequence of global-local processing in 3-month-olds. Infancy, 1, 375-386.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. Cognitive Psychology, 15, 411-458.

Gerken, L.-A. (1996). Phonological and distributional information in syntax acquisition. In J. Morgan & C. Demuth, (Eds.), Signal to Syntax (pp. 411-425). Mahwah, NJ: Erlbaum.

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, 237-265.

Ghim, H., & Eimas, P. (1988). Global and local processing by 3- and 4-month-old infants. Perception and Psychophysics, 43, 165-171.

Gleitman, L., Gleitman, H., Landau, B., & Wanner, E. (1988). Where learning begins: Initial representations for language learning. In F.J. Newmeyer (Ed.) Linguistics: The Cambridge Survey, Vol. 3, Language: Psychological and Biological Aspects (pp. 150-193). New York: Cambridge University Press.

Gleitman, L., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), Language Acquisition: The state of the art (pp. 3-48). Cambridge, UK: Cambridge University Press.

Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P.W., Wright, K., Druss, B., & Kennedy, L.J. (1987). Clauses are perceptual units for young infants. Cognitive Psychology, 24, 252-293.

Johnson, E., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of Memory and Language, 44, 548-567.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.

Jusczyk, P.W. (1998a). Dividing and conquering the linguistic input. In M.C. Gruber, D. Higgins, K. Olson, & T. Wysocki (Eds.), CLS 34, Vol. 2: The Panels (pp. 293-310). Chicago: University of Chicago.

Jusczyk, P.W. (1998b). Using the headturn preference procedure to study language acquisition. In C. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), Advances in Infancy Research, Vol. 12 (pp. 188-204). Stamford, CT.: Ablex.

Jusczyk, P.W., & Aslin, R. N. (1995). Infants' detection of sound patterns of words in fluent speech. Cognitive Psychology, 29, 1-23.

Jusczyk, P.W., Hirsh-Pasek, K., Kemler Nelson, D., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. Cognitive Psychology, 24, 252-293.

Jusczyk, P.W., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research (pp. 91-119). Timonium, MD: York Press.

Jusczyk, P.W., Houston, D., & Newsome, (1999). The beginning of word segmentation in English-learning infants. Cognitive Psychology, 39, 159-207.

Jusczyk, P.W., & Kemler Nelson, D.G. (1996). Syntactic units, prosody and psychological reality during infancy. In J. Morgan & C. Demuth, (Eds.), Signal to Syntax (pp. 389-408). Mahwah, NJ: Erlbaum.

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.s

Klatt, D.H. (1975). Vowel lengthening is syntactically determined in a connected discourse. Journal of Phonetics, 3, 129-140.

Kuhl, P.K., & Miller, J.D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. Perception and Psychophysics, 31, 279-292.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Mandel, D., Juscyzk, P.W., & Kemler Nelson, D.G. (1994). Does sentential prosody help infants organize and remember speech information? Cognition, 53, 155-180.

Mandel, D., Kemler Nelson, D.G., Jusczyk, P.W. (1996). Infants remember the order of words in a spoken sentence. Cognitive Development, 11, 181-196.

Mehler, J., Jusczyk, P.W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 143-178.

Morgan, J.L., Meier, R.P., & Newport E.L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. Cognitive Psychology, 19, 498-550.

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.

Morgan, J.L., Swingley, D., & Miritai, K. (1993). Infants listen longer to speech with extraneous noises inserted at clause boundaries. Paper presented at the Biennial Meeting of the Society for Research in Child Development, New Orleans, LA.

Morse, P.A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. Journal of Experimental Child Psychology, 13, 477-492.

Myers, J., Jusczyk, P.W., Kemler Nelson, D.G., Charles-Luce, J., Woodward, A.L. & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. Journal of Child Language, 23, 1-30.

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.

Neeleman, A., & Reinhart, T. (1998). Scrambling and the PF interface. In W. Gueder and M. Butt (Eds.), Projecting from the Lexicon (pp. 309-353). Stanford: CSLI.

Nespor, M., Guasti, M.-T., & Christophe, A. (1996). Selecting word order: The Rhythmic Activation Principle. In U. Kleinhenz (Ed.), Interfaces in Phonology (pp. 1-26). Berlin: Akademie Verlag.

Nooteboom, S., & Terken, J.M.B. (1982). What makes speakers omit pitch accents? An experiment. Phonetica, 39, 317-336.

Peters, A. (1983). The units of language acquisition. Cambridge: Cambridge University Press.

Pinker, S. (1984). Language learnability and language development. Cambridge, MA: Harvard University Press.

Pinker, S. (1989). Learnability and cognition. Cambridge, MA: MIT Press.

Pierrehumbert, J.B., & Hirschberg, J. (1990). The meaning of intonational contours in interpretation of discourse. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, (Eds.), *Intentions in Communication*, (pp. 271—311). Cambridge, MA: MIT Press.

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. Annual Review of Language Acquisition, 2, 85-115.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. Science, 274, 1926-1928.

Scott, D. (1982). Duration as a cue to the perception of a phrase boundary. Journal of the Acoustical Society of America, 71, 996-1007.

Seidl, A. (2001). Minimal indirect reference: A theory of the Syntax-Phonology interface. New York: Routlege.

Selkirk, E. O. (1996) The prosodic structure of function words. In J. Morgan & C. Demuth, (Eds.), Signal to syntax (pp. 187-213). Mahwah, NJ: Erlbaum.

Shattuck-Hufnagel, S., & Turk, A. (1996). A prosody tutorial for investigators of auditory sentence processing. Journal of Psycholinguistic Research, 25, 193-247.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. Linguistic Inquiry, 30, 219-255.

Werker, J., & Tees. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development, 7, 49-63.

Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. (1992). Segmental durations in the vicinity of prosodic boundaries. Journal of the Acoustical Society of America, 91, 1707-1717.

Figure Captions

Figure 1. Pitch track and waveform for noun phrase (NP) “new watches for men”.

Figure 2. Pitch track and waveform for non-unit (NU) “gnu watches for men”

Figure 3. Pitch track and waveform for noun phrase (NP) “people by the hole”

Figure 4. Pitch track and waveform for non-unit (NU) “people buy the whole”

Figure 5. Mean listening time (in seconds) to passage containing noun phrase (NP) target and non-unit (NU) target for 9-month-old infants in Familiarization Groups 1 and 2

Figure 6. Mean listening time (in seconds) to passage containing noun phrase (NP) target and non-unit (NU) target for 6-month-old infants in Familiarization Groups 1 and 2

Figure 7. Pitch track and waveform for verb phrase (VP) “design telephones”

Figure 8. Pitch track and waveform for non-unit (NU) “design telephones”

Figure 9. Pitch track and waveform for verb phrase (VP) “promise surprises”

Figure 10. Pitch track and waveform for non-unit (NU) “people buy the whole”

Figure 11. Mean listening time (in seconds) to passages containing verb phrase (VP) target and non-unit (NU) target for 6-month-old infants in Familiarization Groups 1 and 2

Table 1

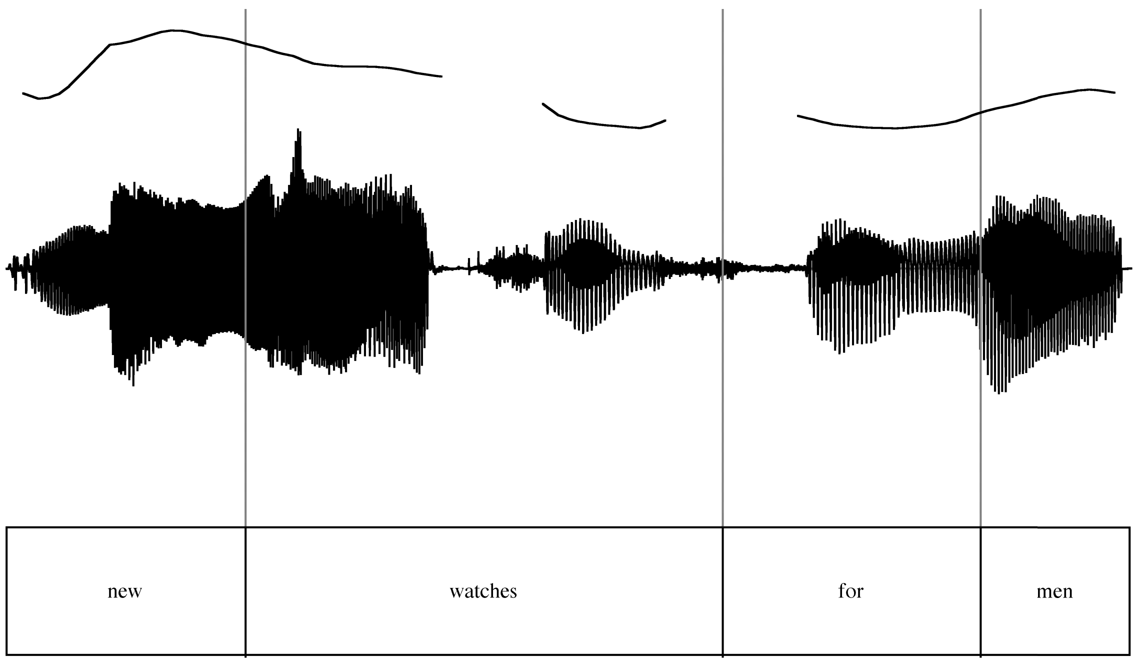
Acoustical measurements of subject/VP boundary in non-unit (NU) sequence compared with same location in noun phrase (NP) sequence in stimuli from Experiments 1 and 2.

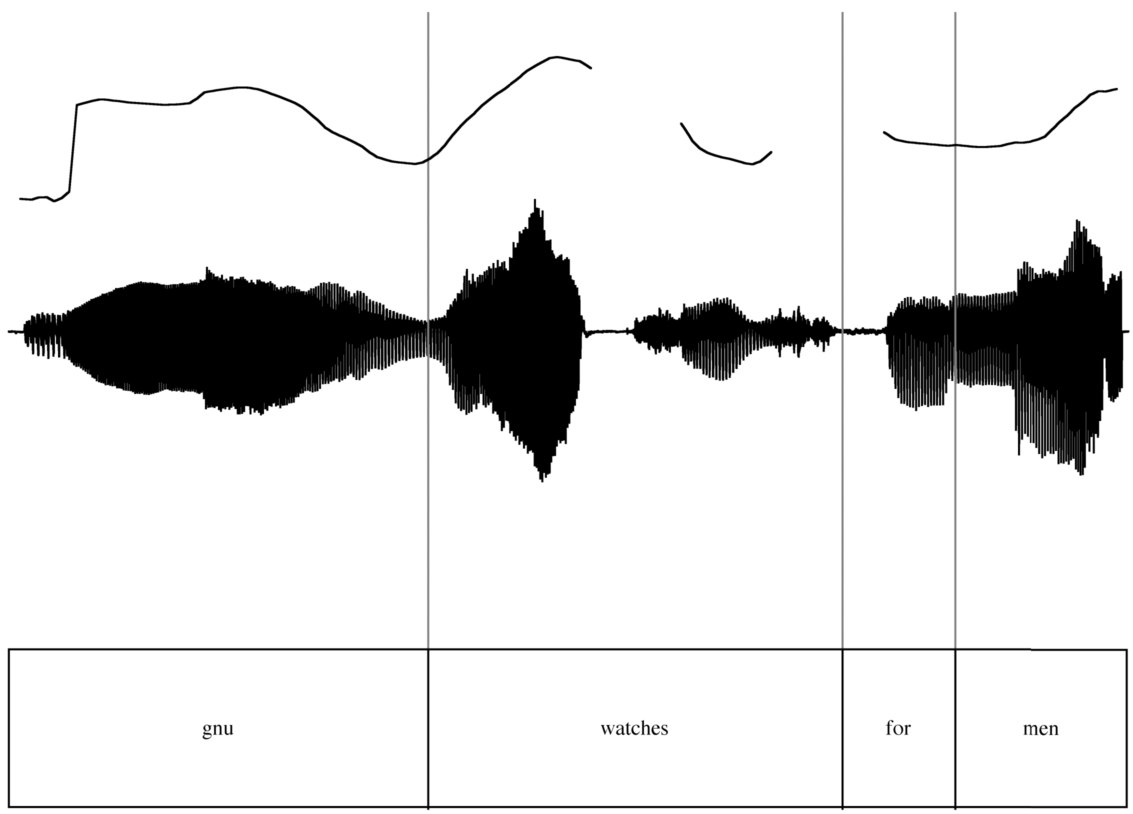
	Group 1 "new/gnu watches for men"		Group 2 "people by/buy the hole/whole"	
Analysis	Noun Phrase	Non-Unit	Noun Phrase	Non-Unit
Final Vowel Length (s)	.089	.229	.074	.092
Pause Duration (s)	0	0	0	.018
High Pitch (Hz)	535.5	414.7	547.5	407.7
High Final Syll (Hz)	535.5	358.8	547.5	308.5
Low Pitch (Hz)	224.7	214.7	203.3	215.8
Low Final Syll (Hz)	317.3	214.7	308.5	305.8
Dif. Fin. Syll (Hz)	218.2	144.1	239.0	23.9
Highest point	new	wa	ple	whole
Lowest point	es/for	gnu	hole	by

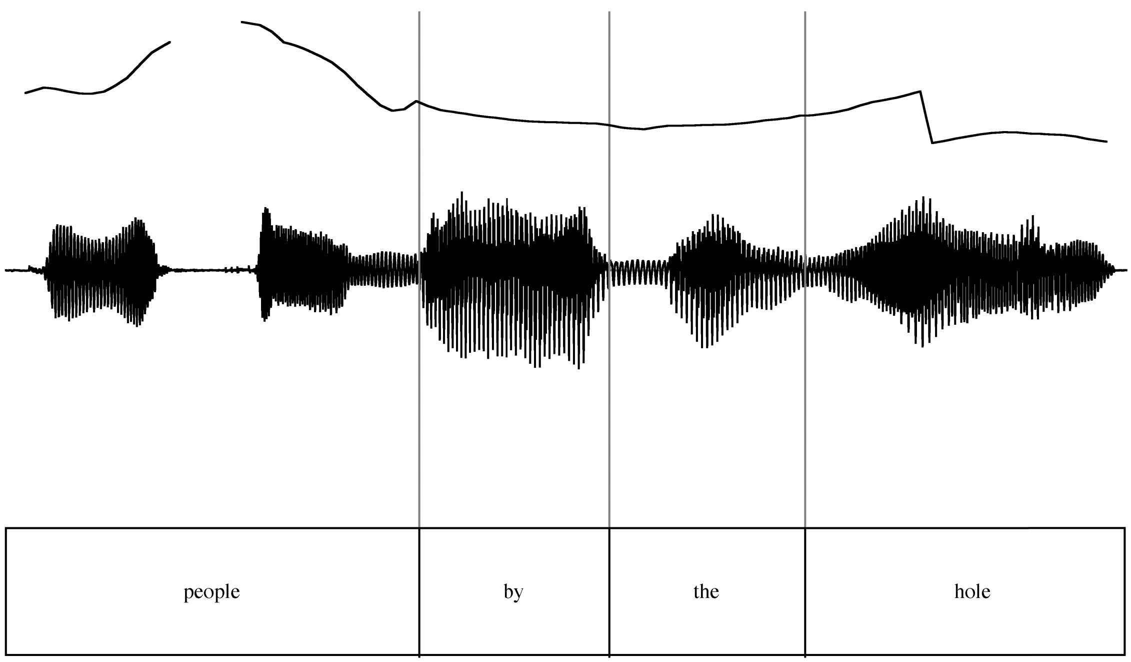
Table 2

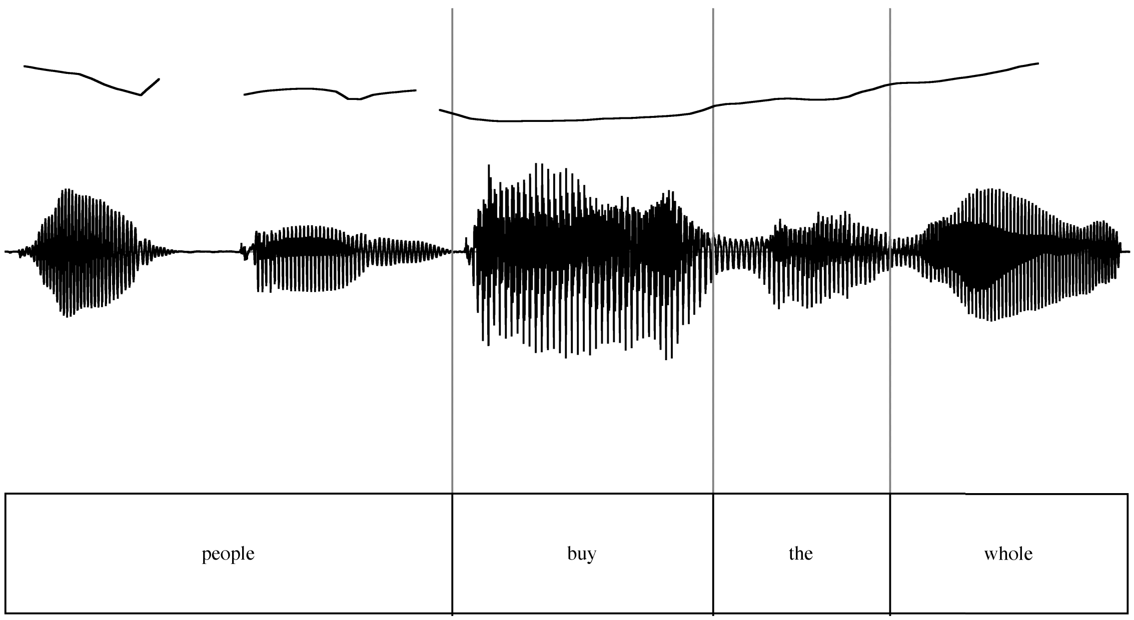
Acoustical measurements of subject/VP boundary in non-unit (NU) sequence compared with same location in verb phrase (VP) sequence in stimuli from Experiment 3.

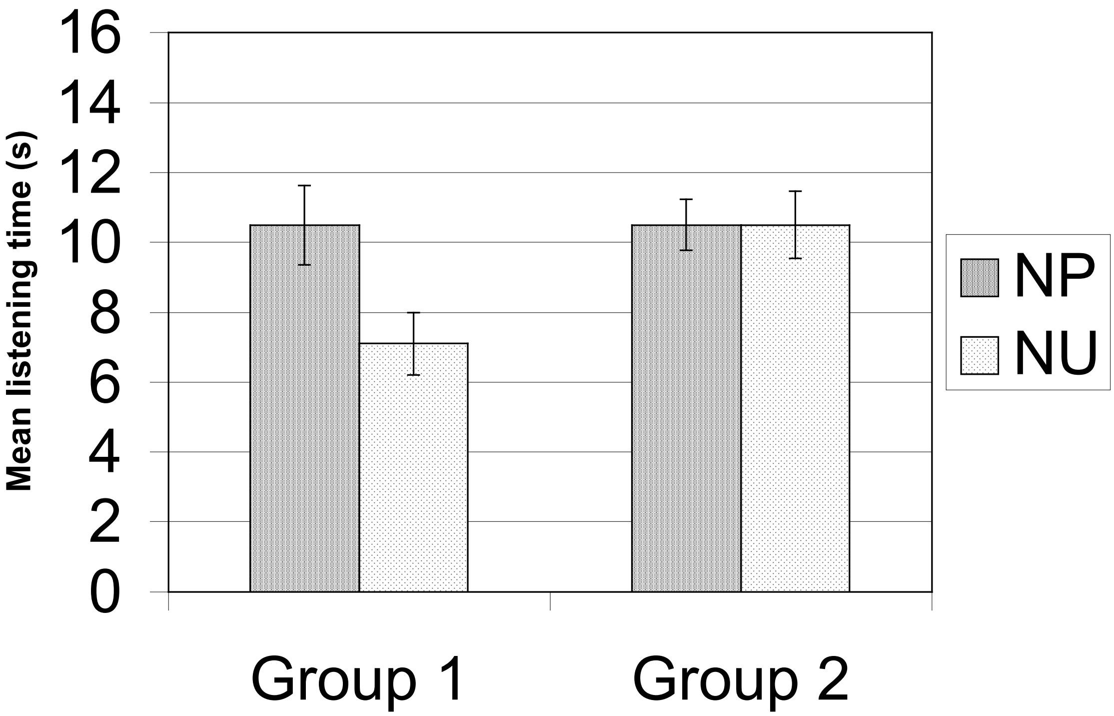
	Group 1 "design telephones"		Group 2 "promise surprises"	
Analysis	Verb Phrase	Non-Unit	Verb Phrase	Non-Unit
Final Vowel Length (s)	.155	.220	.075	.091
Pause Duration (s)	.016	.046	0	.024
High Pitch (Hz)	401.7	559.8	551.2	389.7
High Final Syll (Hz)	269.8	508.9	313.6	309.4
Low Pitch (Hz)	229.7	235.7	285.0	233.2
Low Final Syll (Hz)	248.5	235.7	285.0	245.5
Dif. Fin. Syll (Hz)	21.3	273.2	28.6	63.9
Highest point	tele	tele	pri	pri
Lowest point	phones	sign	sur	sur

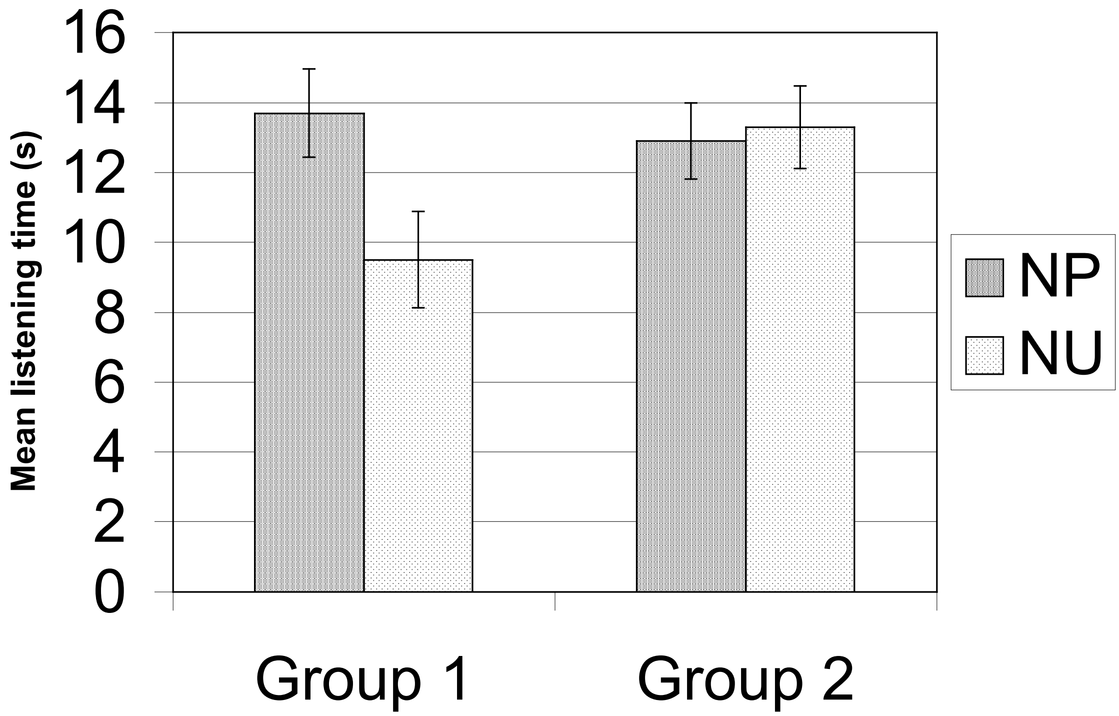

new
watches
for
men


gnu
watches
for
men

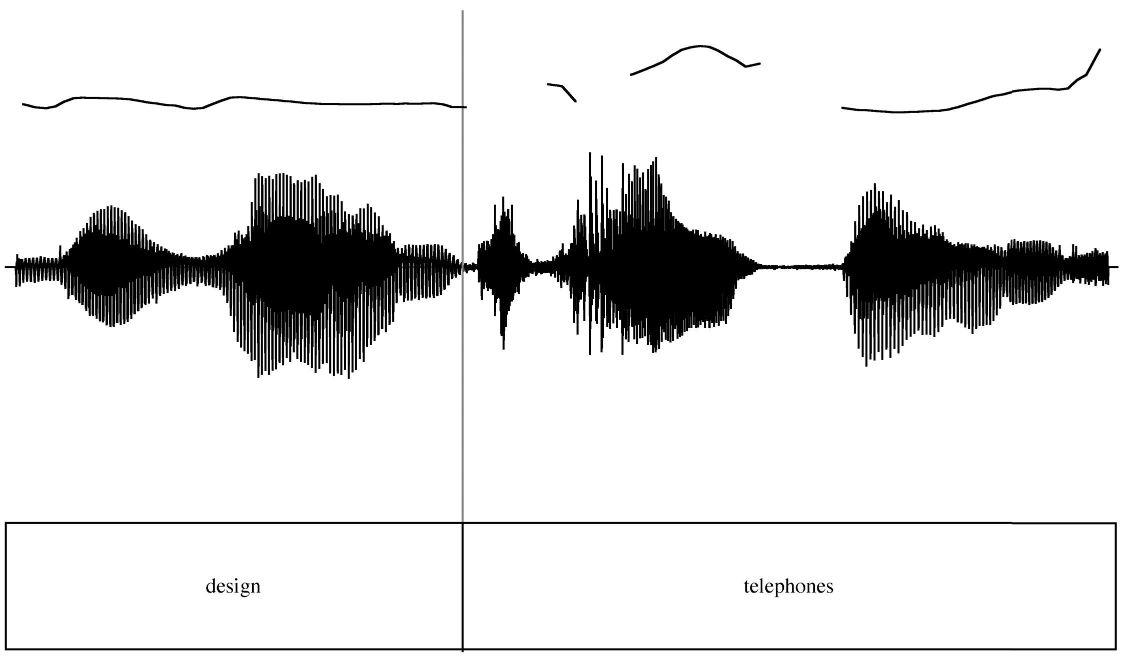

people
by
the
hole

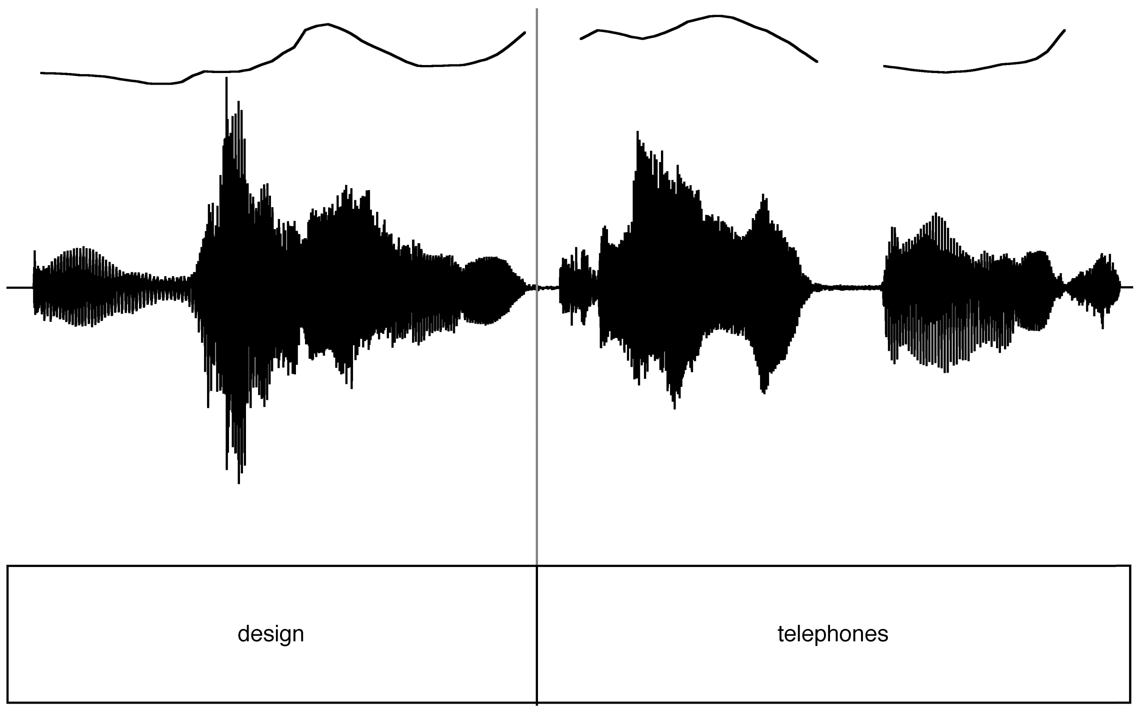

people
buy
the
whole

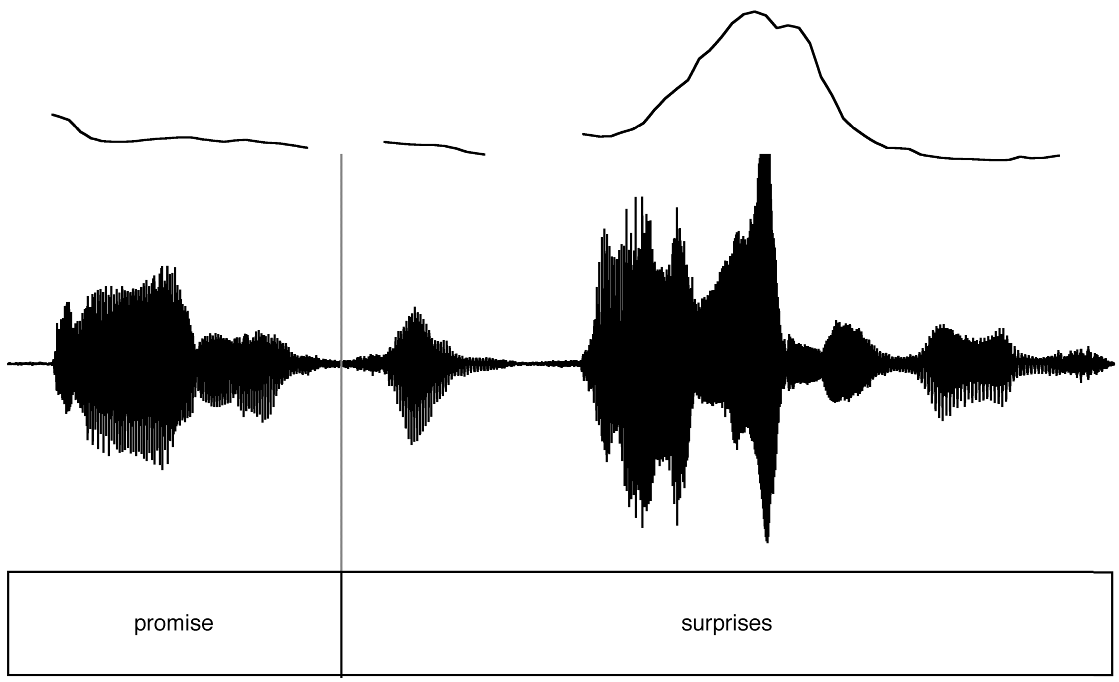

16
14
12
10
8
6
4
2
0
Mean listening time (s)
NP
NU
Group 1
Group 2

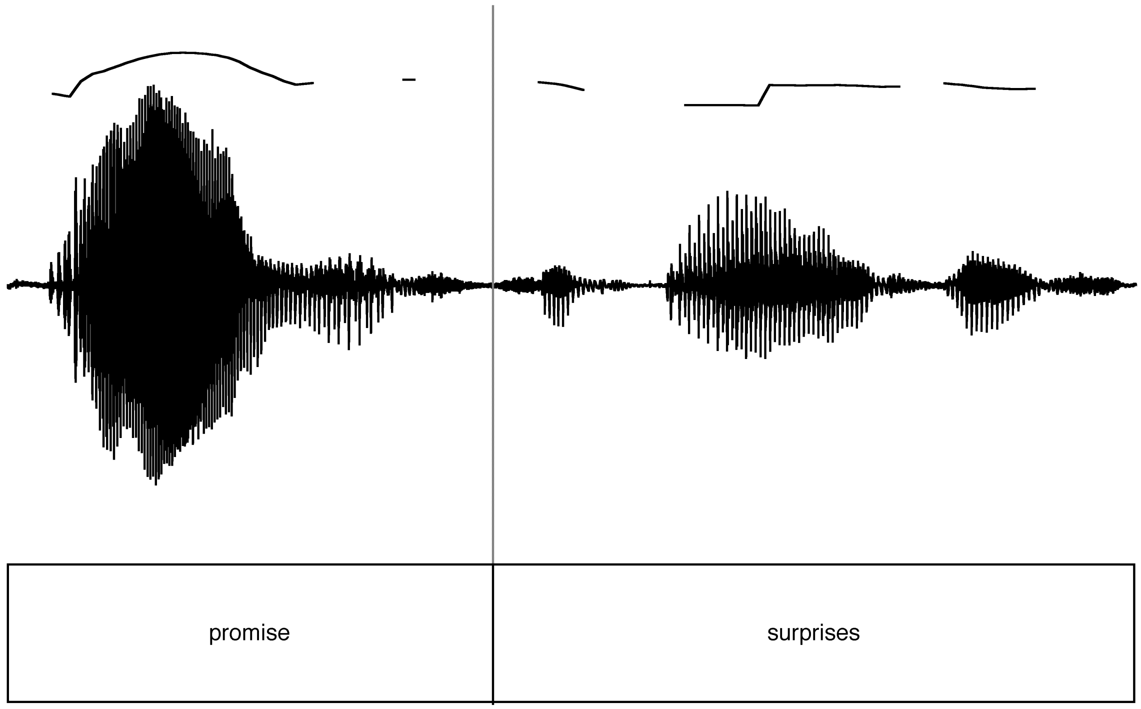


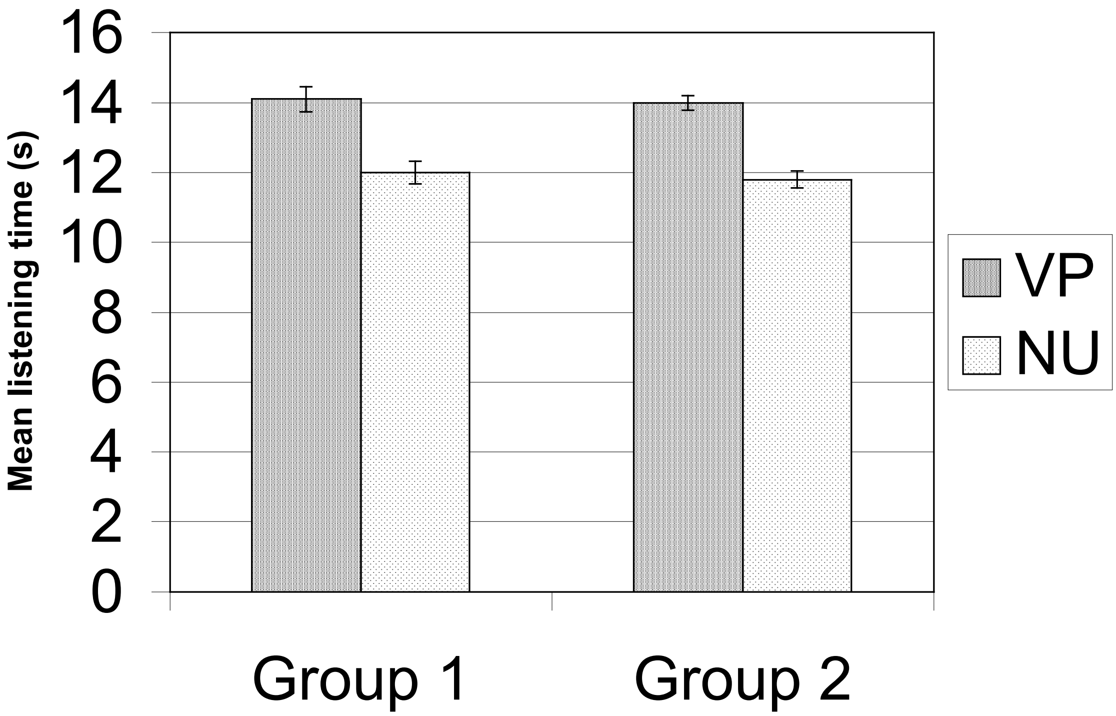
Mean listening time (s)
16
14
12
10
8
6
4
2
0
Group 1
Group 2
NP
NU


design
telephones


design
telephones


promise
surprises


promise
surprises


16
14
12
10
8
6
4
2
0
Mean listening time (s)
Group 1
Group 2
VP
NU