PAPER

Developmental changes in the weighting of prosodic cues

Amanda Seidl and Alejandrina Cristià

Department of Speech, Language and Hearing Sciences, Purdue University, USA

Abstract

Previous research has shown that the weighting of, or attention to, acoustic cues at the level of the segment changes over the course of development (Nittrouer & Miller, 1997; Nittrouer, Manning & Meyer, 1993). In this paper we examined changes over the course of development in weighting of acoustic cues at the suprasegmental level. Specifically, we tested English-learning 4-month-olds' performance on a clause segmentation task when each of three acoustic cues to clausal units was neutralized and contrasted it with performance on a Baseline condition where no cues were manipulated. Comparison with the reported performance of 6-month-olds on the same task (Seidl, 2007) reveals that 4-month-olds weight prosodic cues to clausal boundaries differently than 6-month-olds, relying more heavily on all three correlates of clausal boundaries (pause, pitch and vowel duration) than 6-month-olds do, who rely primarily on pitch. We interpret this as evidence that 4-month-olds use a holistic processing strategy, while 6-month-olds may already be able to attend separately to isolated cues in the input stream and may, furthermore, be able to exploit a language-specific cue weighting. Thus, in a way similar to that in other cognitive domains, infants begin as holistic auditory scene processors and are only later able to process individual auditory cues.

Introduction

Infants as young as 6 months have been shown to be able to use prosodic information to segment clausal units when processing both their native language (Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss & Kennedy, 1987; Nazzi, Kemler Nelson, Jusczyk & Jusczyk, 2000; Seidl, 2007) and an unfamiliar language (Jusczyk, 1989, 2003; Seidl, Johnson, Redman & Brentari, 2004). For example, in a landmark study, Hirsh-Pasek *et al.* (1987) inserted silences in passages of infant-directed speech at either clause or non-clause boundaries and found that 6-month-old infants preferred to listen to passages of speech with interruptions at clause boundaries over passages with interruptions internal to the clause. This preference forms the basis for segmentation of clausal units, an ability which helps infants with tasks such as deciphering their language's syntactic organization (Morgan & Demuth, 1996) and segmenting words by focusing on the edges of clauses (Gout, Christophe & Morgan, 2004; Seidl & Johnson, 2006). Since clause segmentation is a cornerstone in the development of speech perception, we would like to know whether attention paid to each of the relevant cues to clausal units changes as infants' linguistic exposure increases and as a result of development.

Three cues have been highlighted as being strongly linked with clausal boundaries. First, vowels and other syllabic nuclei may be longer when at the end of a clause than in the middle of one (Beckman & Edwards, 1990; Cooper & Paccia-Cooper, 1980; Klatt, 1975; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992), a cue which we will refer to as preboundary lengthening. Second, f_0, or pitch, may signal a clause boundary through pitch resets or certain intonational patterns (Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988).[1] Third, pauses between clauses may be longer than pauses within clauses (Cooper & Paccia-Cooper, 1980; Scott, 1982). These three cues are employed differently across languages, such that speakers may rely on a subset of them more than on the others when signaling or interpreting boundaries. Infants must thus tune into the cue-weighting strategy that will be most useful in their language. However, in order to compute how reliable a cue is as a boundary marker, one must begin by finding the boundary itself, and only then correlate the presence of each cue with the presence of a boundary.

Previous language acquisition research suggests a time course for the segmentation of grammatical units such as clauses and phrases (e.g. Jusczyk, 1997; Soderstrom, Seidl, Kemler Nelson & Jusczyk, 2003; Nazzi, Jusczyk & Johnson, 2000). It has also been shown that infants attend to many isolated aspects of prosody at multiple ages (rhythm; Ramus, 2002; and various aspects of pitch

---

[1] The relative importance of pitch resets as compared to intonational patterns in English listeners has not been investigated. In principle, however, either may act as a key cue delineating a prosodic boundary.

Address for correspondence: Amanda Seidl, Department of Speech, Language and Hearing Sciences, Purdue University, W. Lafayette, IN 47907, USA; e-mail: aseidl@purdue.edu

and pitch alignment; Demany, McKenzie & Vurpillot, 1977; Nazzi, Floccia & Bertoncini, 1998; Bosch & Sebastián-Gallés, 1997). However, in these studies the three main acoustic cues to clauses, namely pitch, preboundary lengthening and pause, were not manipulated separately. Therefore, although this previous research demonstrates that infants are sensitive to prosodic cues, it does not show whether these cues are differentially weighted and, if so, how infants converge on a cue-weighting strategy.

One recent study in our lab has begun to address these questions. Seidl (2007) found that English-learning 6-month-old infants weight pitch more heavily than either pause or preboundary-length cues to clausal boundaries, although we found that converging cues of pitch and pause or pitch and preboundary length were necessary for infants to successfully segment clausal units. Specifically, in a series of experiments using the Headturn Preference procedure, the acoustic cues that infants might use to segment clausal units from continuous speech were systematically neutralized. It was found that performance on this clausal segmentation task was severely degraded when pitch was neutralized, but was not disturbed by neutralization of pause or preboundary length alone. In this study, 6-month-old infants were familiarized with a sequence of words as a complete clause and the same sequence of words spanning a clause boundary (a 'non-clause'). In the Baseline condition, pitch, preboundary length, and pause length were significantly different across these two realizations of the sequence. During testing, infants successfully discriminated between a passage that contained the familiar clause and one including the familiar non-clause. The following experiments manipulated each of the acoustic cues so that the familiarization stimuli did not differ in that particular cue. For example, pitch either was flattened to a declining contour in both the clause and the non-clause, or the intonation patterns of clause and non-clause were swapped (so that the string excised from the end of one clause and beginning of the following exhibited the boundary markers of lengthened vowel and a pause in the middle of the string, but the overall pitch pattern was the same as that of a clause). Infants in these experiments did not exhibit the same pattern of discrimination when pitch was manipulated, but their performance did not suffer when either of the other individual cues was neutralized. Therefore, 6-month-old infants rely heavily on pitch, since neutralization of intonation as a distinguishing factor between clause and non-clause rendered the segmentation task impossible. Likewise, pause and preboundary length differences between clause and non-clause were not necessary for infants to succeed in this study, showing that they likely do not rely on these cues for segmentation.

Further research suggests that these results are not due to a universal sensitivity to pitch cues over other cues, but rather to the English-learning infants knowing that pitch cues are more reliable in their ambient language. The basis for this conclusion lies in the comparison of English and Dutch infants and adults. Unlike English-learning infants, Dutch-hearing infants rely heavily on pause in a clause segmentation task (Johnson & Seidl, 2005, in press). This may be related to the fact that speakers in different linguistic environments weight acoustic cues to prosodic boundaries differently. While there is a paucity of work on how adults weight cues to prosodic boundaries, the crosslinguistic work done strongly suggests such diversity. For example, English-speaking adults seem to rely more heavily on preboundary length than on pause in detecting clausal boundaries (Aasland & Baum, 2003), and likewise Russian adults do not weight pause very heavily (Volskaya, 2003). Conversely, Dutch (Sanderman & Collier, 1997) and Swedish speakers (Horne, Strangert & Heldner, 1995; House, Hermes & Beaugendre, 1998) do rely heavily on pause duration for prosodic boundary judgments. If, as this evidence suggests, the 6-month-olds' ability to attend to one prosodic cue over another may be an effect of language exposure, we would like to find out whether there is an earlier stage of development wherein infants possess a different cue-weighting and what drives this different cue-weighting strategy.

In order to test whether developmental changes occurred for the weighting of acoustic cues linked to clausal boundaries between 4 and 6 months of age, we conducted an experiment with a Baseline and three experimental conditions corresponding to the manipulation of each of three prosodic cues to clausal units. The Baseline condition probed whether 4-month-old infants treat clauses and non-clauses differently in a segmentation task when all cues are present. If they do, this would constitute evidence that infants are already paying attention to grammatical boundaries, which raises the question of which prosodic cues they are using to effect the discrimination. The three prosodic cues we manipulated are the three prosodic correlates of clausal units that are most well documented in the literature: preboundary length, pitch, and pause. Thus, we tested infants in a Pause neutralization condition, which probes whether English-learning 4-month-olds weight pause more heavily than 6-month-olds do. A Pitch neutralization condition probes whether pitch is as salient at 4 as it is at 6 months. Finally, a Preboundary length neutralization condition attempts to determine whether 4-month-olds rely upon preboundary length cues more than 6-month-olds do.

Experiment

In this experiment we used the Headturn Preference procedure (Jusczyk & Aslin, 1995; Kemler Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989) to explore infants' weighting of prosodic cues to clausal boundaries. There were three manipulation conditions corresponding to the neutralization of each cue, Pitch, Pause, and Preboundary Length, as well as a Baseline condition in

Table 1 *Familiarization conditions and acoustic cue manipulations*

Condition	Subcondition	Acoustic cue manipulation
Baseline		None
Preboundary length	Short	Vowel before clause boundary shortened to within-clause duration
	Long	Vowel within clause lengthened to before-clause boundary duration
Pause	Short	Pauses between clauses shortened to within-clause pause duration
	Long	Pauses within clauses lengthened to between-clause pause duration
Pitch		Pitch flattened to declining contour in both clause and non-clause stimuli

which all cues were present. In each condition, infants were familiarized with a string of words, and were later tested on passages that presented this string of words either as a clause or a non-clause (that is, the string of words spanned a clause boundary). Successful segmentation in each condition is shown by differential listening to each of these test passages. Thus, if infants failed to attend differently to these test passages in one of the manipulation conditions, as compared to Baseline, then we assume that this cue is a necessary one.

Method

Participants

We tested 96 4-month-old normal hearing infants from a Midwestern college town with a mean age of 4.2 months (range = 3.5 mos to 4.5 mos) with 24 infants in each of the four main conditions (Baseline, Pause neutralization, Pitch flattening, and Preboundary length neutralization). Forty babies were excluded from the final analysis, 36 for fussing or crying and four for experimenter error. Within the Pause and Preboundary length manipulation groups there were two subconditions, such that 12 of the infants were run on stimuli neutralized by lengthening non-boundary pause or vowel durations to the between-clauses duration and 12 were run on stimuli neutralized by shortening boundary pauses and vowel durations to the within-clause length.

Design

This experiment used a between-subjects design and the same procedure as Nazzi, Kemler Nelson *et al.* (2000) and Seidl (2007). In our experiment, infants were familiarized with a word sequence spoken in infant-directed speech as both a clause and a word sequence that spanned a clause boundary (a non-clause). Afterwards, they were tested with passages containing the clause or the non-clause. Infants, as in Nazzi, Kemler Nelson *et al.* (2000) and Seidl (2007), should be able to differentiate the recurrence of a familiarized word sequence when it constitutes a well-formed unit from when it occurs as a non-unit. Half of the infants in this study were familiarized with the word sequence *rabbits eat leafy vegetables* and the other half were familiarized with the word sequence *leafy vegetables taste so good* spoken both as a clause

and as a non-clause. The familiarization clauses are shown represented in italics and the familiarization non-clausal sequences are represented in boldface type below.

1. John doesn't know what **rabbits eat**. *Leafy vegetables taste so good*. They don't cost much either.
2. Many animals prefer some things. *Rabbits eat **leafy vegetables***. **Taste so good** is rarely encountered.

If infants are able to find clausal units using acoustic cues, we predicted that infants should show a significant preference for either test item and that this would be found for both groups of infants in the crossed design (i.e. it is not a stimulus effect). One-quarter of these infants were familiarized with stimuli corresponding to each of the following conditions: Baseline, Pause neutralized, Pitch neutralized and Preboundary length neutralized. Then all infants were tested on the same stimuli. The logic of this experiment is the same as Seidl (2007): If we manipulated a necessary cue, then infants should show no preference for clause or non-clause in test.

Familiarization conditions are schematized in Table 1.

Stimuli

As shown in Table 2, there are major acoustic differences between the clause and non-clause stimuli.[2] Syllabic nuclei at the clause boundary are approximately twice as long as those within the clause, pauses at the syntactic clause boundaries are 8–10 times longer than pauses between the same word sequence within the clauses, and pitch resets/discontinuities are found at clause boundaries and not within clauses. Thus there were ample acoustic cues to boundaries in the baseline stimuli. Waveforms, tone analysis (Beckman & Hirschberg, 1994; Venditti, 1995) and pitchtracks aligned with the text of the speech from a sample set of Baseline familiarization stimuli, as well as details on the relevant manipulations, are shown in Figure 1.

---

[2] Pitch change was calculated by measuring pitch at the end of the last voiced segment before the putative boundary and at the beginning of the following voiced segment. For instance, we measured pitch in /i/ of 'eat' and /l/ in 'leafy' to determine the pitch change for the Rabbits stimuli. Sample stimuli are available at: http://web.ics.purdue.edu/~aseidl/AmandaSeidl.html.

Table 2 *Acoustic correlates of syntactic boundaries present in baseline test stimuli*

	Leafy group		Rabbits group	
Acoustic correlate	within clause	clause boundary	within clause	clause boundary
Word sequence measured	[. . . vegetables taste . . .]	[. . . vegetables] [taste . . .]	[. . . eat leafy . . .]	[. . . eat] [leafy . . .]
Preboundary length	.12 s	.21 s	.12 s	.24 s
Pause duration	.05 s	.38 s	.03 s	.3 s
Change of pitch	68 Hz	233 Hz	−47 Hz	244 Hz
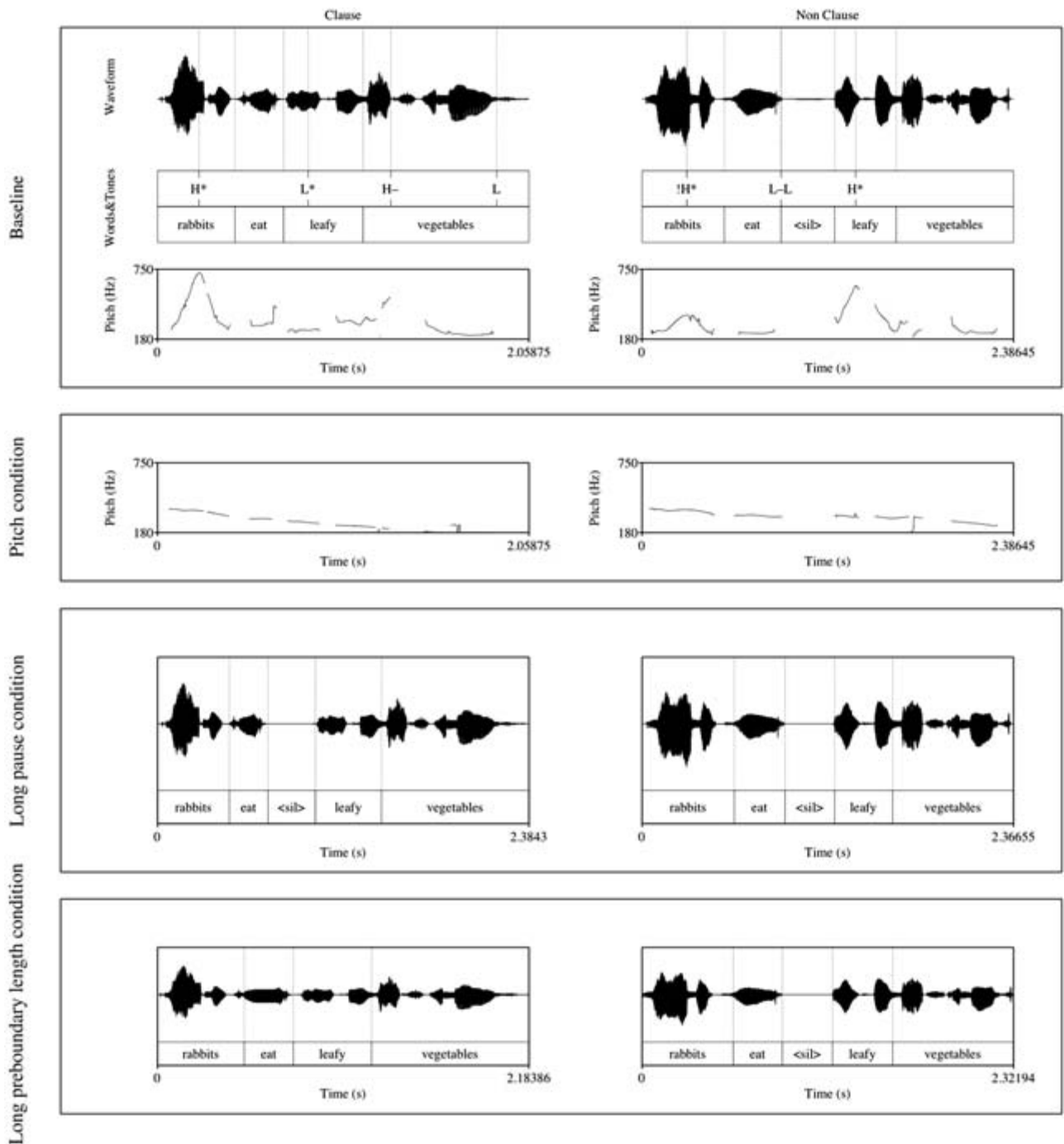
Figure 1 *The left panel shows stimuli classified as 'clause' and the right panel stimuli used for the 'non-clause'. Waveform, tone analysis, text of the speech and pitchtrack are given for the Baseline condition at the top. Immediately below is the modified pitchtrack used for the Neutralized Pitch condition, followed by the Neutralized Pause-long condition, and the Neutralized Preboundary Lengthening-long condition at the bottom. In the latter three, only the relevant acoustic cue was neutralized, while all other cues remained unmodified and equal to those in the Baseline.*

As mentioned, there were four different familiarization conditions. In the Baseline condition all prosodic cues were present and unmodified. There were also Neutralized Pause, Neutralized Preboundary lengthening, and Neutralized Pitch conditions. Modifications to the stimuli for each condition were done using Praat (Boersma & Weenik, 2005). Familiarization stimuli for modified conditions used the same sequences as Baseline, but acoustic cues were modified. To insure that all familiarization stimuli were comparable in naturalness, both clause and non-clause stimuli were modified for each manipulation. On the other hand, the test stimuli were identical across all conditions. In the Neutralized Pause condition, there were four sets of familiarization stimuli (Leafy-Short, Leafy-Long, Rabbits-Short and Rabbits-Long), two with pauses neutralized by lengthening the shorter pause (Long), and two with pauses neutralized by deleting the longer pause (Short) in the same familiarization passages that were used in the Baseline condition. The Neutralized Preboundary Length condition had the same design. In the Short condition, the durations of preboundary lengthened vowels were halved to make them equal in length to the vowels that were not before a clausal boundary. In the Long condition the non-preboundary vowel and syllabic consonant [l] were doubled in length to make them equal in length to the syllables that appeared before a boundary. Finally, in the Neutralized Pitch condition, pitch contours in both the clause and the non-clause sound files were flattened to a generalized declining contour. We chose to do this modification, rather than switching the pitch as was done in one of the experiments in Seidl (2007), because there were no significant differences between these two manipulations in the previous set of studies with 6-month-old infants and we thought that this was an easier task since it did not employ clearly conflicting cues.

It may be argued that any of the modifications that we performed on our stimuli detract from their naturalness. For example, it is not natural to hear a clause where the vowel preceding the boundary is short, given that English speakers always produce lengthened vowels before a boundary. However, if the modified aspect is not relevant perceptually, its modification should not impact task performance. An example of a modification that does not produce natural speech but did not detract from infants' performance is the common method of low-pass filtering that has been used to test infants' ability to discriminate utterances in different languages (e.g. Bosch & Sebastián-Gallés, 1997). On the other hand, if the acoustic correlate that has been tampered with is relevant, that is, if infants use it as a cue to effect a discrimination, then we may expect degraded performance.

Apparatus and reliability

A Macintosh G4 computer controlled the presentation of the speech samples and recorded an observer's coding of the infants' looking preferences by a button box. The audio signal was fed through an audio amplifier to Cambridge Soundworks Ensemble II speakers. The testing booth consisted of a three-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 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 allowed a secondary observer, such as a second parent or student observer, to view the procedure. A cushioned 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 speakers was situated behind the two side panels, located directly behind the red lights. The computer, response box, and other equipment were located behind the front panel, out of sight of the infant.

Subsequent to each experiment in this paper, videotapes of 19 infants, 4–5 from each condition or 20% of the total sample, were randomly selected to be recoded offline for reliability by a second observer. Intercoder reliability was 98.4 between the live and videotape observers for each condition.

Procedure

A modified version of the Headturn Preference Procedure with both a test and familiarization phase (Jusczyk & Aslin, 1995; Kemler Nelson *et al.*, 1989) was used. Each infant was seated on the caregiver's lap on the chair in the middle of the three-sided booth. The experimenter was situated behind the front panel, 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 side light and specific auditory stimulus were made randomly via computer program. Both the experimenter and the caregiver wore tight-fitting headphones (Peltor Aviation Headset 7050) 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 green light flashing to attract the infant's attention to the center. When the infant looked to the center light, this light would extinguish, and one of the two side red lights would begin to flash. When the infant oriented to the side light, one of the auditory test strings would play repeatedly. This continued until either the infant looked more than 30 degrees 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 seconds and then looked back again, the trial continued, but the amount of time spent looking away was not counted in the overall tally.

The experiment consisted of two phases: Familiarization and test. During familiarization, two versions of the same word sequence (clause and non-clause) played alternately until the infant accrued 30 seconds of orientation time to each sequence. If the infant looked away for more than 2 seconds during the playing of one of these word sequences it would play until the end of the word sequence. This was done to avoid familiarizing the infant with something that was neither a non-clause nor a clause. During the test phase infants heard passages that contained either familiarized word sequence. Three blocks of randomly ordered test trials were presented, with a total of four trials per block. Two out of four of these trials contained clausal test items and the other two contained non-clausal test items. Each test trial contained a test passage which would repeat a maximum of three times with an intervening pause of 1 s. The dependent measure was the average looking time across trials to each stimulus type.

Results

Before analyzing the test results, we ran some preliminary analyses on the familiarization stage. Even though all infants were familiarized until they reached a cumulative minimum of 30 s of orientation time to each of the clause and non-clause stimuli, they could have listened slightly longer to one of these sequences in familiarization since they were randomly repeated until the infant had accumulated 30 s orientation time to each sequence. Thus it is possible that the infants could have preferred one version over the other simply because they heard it more in familiarization. However, analyses of the infants' overall average orientation times during familiarization revealed that infants oriented an average of 24.29 s (*SD* = 7.63 s) to the clausal version and 24.19 s (*SD* = 7.37) to the non-clause version of the familiarization stimuli.[3] This difference was not reliable, $t(23) = .13$. To eliminate the possibility that Baseline familiarization orientation times may have been different than those in the other conditions we also ran an ANOVA with the difference scores between clause and non-clause orientation time during familiarization as the dependent measure and Condition and Group as factors. This revealed no main effects of Group, $F(1, 95) = 1.11, p = .29$, Condition, $F(3, 95) = 1.37$, $p = .25$, and no interaction, $F(3, 95) = 1.96, p = .12$.

[3] Note that these may be averages across multiple trials since some infants required multiple trials of the same stimulus in order to accumulate the 30 s criterion while others did not.
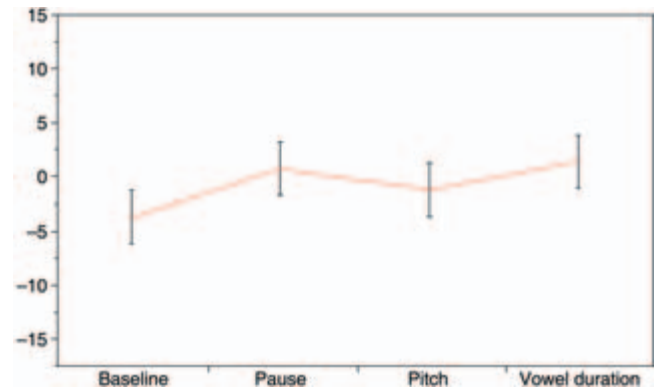
Figure 2 *Mean looking time differences to clause and non-clause in seconds with standard error bars.*

Further, given that there were two subconditions in two of our manipulation conditions, we wondered whether the subconditions were different from one another. For this reason, we compared the listening time difference scores (clause–non-clause) in the neutralized short subconditions to the long subconditions in each of Pause and Preboundary length. Results of each of these two *t*-tests revealed no significant difference in listening times between the two subconditions for the long and short subcondition of the Preboundary Length condition ($t(23) = .09$) or the long or short subconditions of the Pause condition ($t(23) = .31$). Thus in all subsequent analyses we collapse the Long and Short subconditions in Pause, and the Long and Short subconditions of Preboundary length.

Next, we ran a repeated measures ANOVA with Condition (Baseline, Pause, Preboundary length, Pitch) and Group (Leafy, Rabbits) as between-subjects variables and Listening time (Clause, Non-Clause) as our within-subjects variable. This revealed no main effect of Listening time, $F(1, 88) = .017$, Condition, $F(3, 88) = .20$ or Group, $F(1, 88) = .41$. There was a significant interaction between Listening time and Condition, $F(3, 88) = 3.42$, $p < .020$, but there were no interactions between Listening time and Group, $F(1, 88) = 1.4, p = .23$ or Listening time, Group and Condition, $F(3, 88) = 1.28, p = .283$ (see Figure 2).

Given the interaction between Condition and Listening time, we would like to know which conditions were different. We performed a Dunnett's test on the difference scores (clause–non-clause) for each condition with Baseline specified as the control group. This post-hoc test revealed that while the Pitch condition did not differ significantly from Baseline ($LSD = -1.67, p = .34$), both Preboundary length ($LSD = .898, p = .01$) and Pause ($LSD = .226, p = .04$) did.

T-tests comparing listening times to clauses and non-clauses within the Baseline condition revealed that infants listened significantly longer to non-clauses than clauses ($t(23) = 3.18, p < .004$). These same tests within Pause ($t(23) = .42$), Preboundary length ($t(23) = 1.14, p =$

**Table 3** *Mean listening times and standard deviations for each condition*

Condition	Clause	Non-clause
Baseline	15.1 (6.67)	18.8 (7.01)
Pitch	15.7 (7.57)	16.9 (8.04)
Pause	17.2 (5.99)	16.6 (7.1)
Preboundary length	18.4 (7.17)	17.08 (6.87)

Figure 3 *Developmental shifts from holistic to analytic processing.*

.26) and Pitch ($t(23) = 1.23$, $p = .23$) conditions revealed no such differences. Means and standard deviations from all four conditions are shown in Table 3.

Discussion

The results of this experiment build on those of Jusczyk (1989, 2003) by suggesting that not only are 4-month-old infants sensitive to clausal prosody, but they already use prosody to discriminate between clauses and non-clauses in a segmentation task. Further, results reveal that English-learning infants at this age are sensitive to manipulation of all three of the cues, pitch, pause and preboundary length, such that neutralization of each of these cues causes infants to be unable to differentiate the clause from the non-clause at test.

The results of this experiment thus suggest that 4-month-olds apply a holistic strategy in segmenting clauses by attending to all the cues, such that a significant difference between clause and non-clause for each and every cue is a necessary condition for successful segmentation. This is a reasonable strategy since, at this age, infants are only starting to pay attention to the edges of clausal units, and thus they initially recruit all correlates, so that removal of any of the cues degrades segmentation compared to a Baseline condition where all cues are present. We will refer to this strategy as a *holistic sensitivity strategy*. There is some evidence in favor of such a holistic sensitivity in other domains of infants' speech perception, as well as other cognitive domains. With respect to speech, the development of infants' phonological knowledge seems to progress from general prosodic patterns such as rhythm (Nazzi, Bertoncini & Mehler, 1998; Ramus, 2002) at birth, to the clausal level at 4–6 months (Hirsh-Pasek *et al.*, 1987; Nazzi, Kemler Nelson *et al.*, 2000; Seidl, 2007), to the word level at 7.5 months (Jusczyk & Aslin, 1995), to the level of the sound around 10–12 months (Werker & Tees, 1984), and plausibly tuning into subsegmental characteristics throughout childhood (Hazan & Barrett, 2000). Within each of these domains, the same pattern emerges. Thus, infants are only able to discriminate languages with different rhythmic patterns when younger than 3 months (Christophe & Morton, 1998; Nazzi, Bertoncini & Mehler, 1998), but are able to pick up on even dialectal variation by 5 months of age (Bosch & Sebastián-Gallés, 1997; Nazzi, Jusczyk & Johnson, 2000). In the word segmentation domain, infants initially attend to all
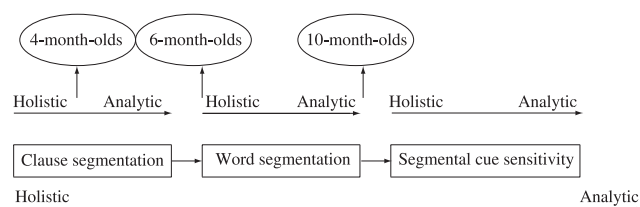
information in the speech signal, being unable to reliably abstract from voice quality until 10.5 months of age (Houston & Jusczyk, 2000; Jusczyk, Houston & Newsome, 1999), nearly 5 months after they have begun to segment words from the stream of speech (Bortfeld, Morgan, Golinkoff & Rathbun, 2005). Likewise, infants are able to distinguish many speech sound contrasts that are not present in their language around 7 months of age, but they tune into their native language's segmental contrasts by their first birthday (Werker & Tees, 1984). In a similar fashion, young children learning the subsegmental cues in their language preferentially attend to larger, syllable-sized units over smaller, segment-based ones even when labelling segments (Nittrouer & Miller, 1997; Nittrouer *et al.*, 1993). These authors suggest that children pay more attention to larger units because they contain dynamic cues spanning longer periods of time. For example, in the processing of place of articulation in fricatives, children favor the dynamic cue of formant transition into a vowel over static and short cues, such as the frequency band of the frication noise, but adults do not. This suggests that children must undergo a 'Developmental Weighting Shift' to reach adult-like cue-weighting, so that they move from larger, more global cues, to smaller, local ones (although see Mayo & Turk, 2004, for a different point of view). In summary, the development of speech perception appears to proceed from larger wholes to smaller parts (Jusczyk, 1997). A visual representation of this process for the first year of development is suggested in Figure 3.

Moreover, this holistic strategy is similar to strategies seen in the development of several other cognitive domains. In the visual domain, a global precedence effect has been documented with infants as young as 3 months; that is, infants exhibit a bias to first attend to global characteristics and only later to process local features in visual discrimination tasks (Colombo, Mitchell, Coldren & Freeseman, 1991; Freeseman, Colombo & Coldren, 1993). Young infants are also able to integrate displays on the basis of lightness of color, and their holistic processing of such displays leads them to generalize to displays that preserve the general shape (for instance, from rows of colored squares to horizontal stripes; Quinn, Burke & Rush, 1993). Four-month-old infants process inverted faces holistically unlike 7-month-old infants, children and adults, who all process inverted faces locally (Cashon & Cohen, 2004). Further,

the tendency to process more general characteristics before more specific ones is also obvious in cross-modal categorization, as shown in the work of Bahrick (Bahrick, 2002) and Quinn (Quinn & Johnson, 2000). For example, Hernandez-Reif and Bahrick (2001) tested categorization of objects presented haptically and visually with respect to amodal relations (those that transcend a single modality, such as shape and size, both perceivable through tactile and visual means) and to arbitrary, modality-specific relations (such as shape and color, the latter being perceivable only visually). They found that 4-month-old infants are able to learn only the amodal relations in a short period of time, whereas 6-month-olds succeed in learning both kinds of relations in the same amount of time. Hernandez-Reif and Bahrick (2001) interpreted these results as congruent with the proposal that 'detection of global (amodal) relations precedes and constrains detection of nested relations (e.g. arbitrary, modality-specific relations)' (p. 53).

Although such a holistic to local processing change might occur in other cognitive domains, it may have been expected that infants would follow a different strategy specifically in clause segmentation, a more linguistically based strategy, if they had been guided by a Universal Grammar. This linguistic weighting would follow from work on child language acquisition that is based on a continuity assumption (Crain, 1991; Pinker, 1989, 1991), which holds that the infant's emerging linguistic system ought to be explained in terms of the same abstract linguistic principles and representations invoked in accounts of adult grammars. Infants guided by a rich Universal Grammar should attend to cues that are linguistically unmarked, that is, more predictable and stable across languages. For example, although pause serves many pragmatic and paralinguistic functions (Goldman-Eisler, 1972), it is correlated with boundaries in all the languages of the world and its only syntactic or prosodic function is as a boundary marker. As a consequence, infants following a linguistic bias may attend primarily to pause in a clause segmentation task. In contrast, both pitch and preboundary length may serve more than one linguistic function across languages. Specifically, they may be recruited to signal lexical contrasts and even segmental contrasts. In addition, preboundary lengthening, unlike pitch, appears to be a universal marker of clausal boundaries, present even in languages with contrastive vowel length (Hayes, 1990; Vaissiere, 1983). In contrast, not only does the use of pitch as a grammatical boundary marker vary widely across languages (Hirst & Di Cristo, 1998) (e.g. with pitch peak alignment varying even within dialects; Bruce & Gårding, 1978; Smiljanic, 2006), but also pitch often does double duty serving other linguistic purposes, such as morphological and lexical contrasts. If infants follow a universal algorithm which searches for large pitch changes to find clauses, they would correctly locate boundaries marked by pitch resets at phrase edges. By the same token, however, infants would

often be misled in tonal languages wherever a low lexical tone precedes a high lexical tone. From a universal linguistic perspective, that represents a serious problem, given that it is estimated that up to 70% of the world's languages are tonal (Yip, 2002). As a consequence, a strategy based on universal grammar which guides infants to clausal boundaries by forcing them to attend to the nearly universal cues of pause and preboundary lengthening over pitch cues would, in theory, greatly aid acquisition.

However, the pattern of results found in this study is not compatible with the linguistic weighting hypothesis. Young infants do not seem to be guided by Universal Grammar, weighting universal cues more heavily than language-specific ones, but appear to weight all cues equally. Thus, this reliance on all three cues seems to follow most directly from a holistic sensitivity, and may suggest a general learning mechanism, rather than a primarily linguistic one.

Using the exact same design as Seidl (2007) further allows us to compare the effects of each cue manipulation in 4- and 6-month-olds' segmentation and to gain an insight into how infants converge in their cue-weighting strategy at each age. A major finding in the present study is the dissimilarity in the cue-weighting of 4- and 6-month-olds. While 6-month-olds are sensitive to preboundary lengthening and pause only as convergent cues with pitch, 4-month-olds' performance is significantly degraded by the absence of each and every cue. Moreover, the increased weight English-learning 6-month-olds allot to pitch cues may be explained through their increased language exposure. By focusing on boundaries, and on strong cues, with sufficient exposure infants might learn that pitch is a reliable cue to grammatical boundaries. The same strategy would lead infants to believe pause to be unreliable as marker of boundaries in their ambient language, since in English pause serves primarily non-syntactic functions (Goldman-Eisler, 1972). In other words, although pause is a correlate of syntactic boundaries (whenever there is a boundary, there is a reliably longer pause), it is not a cue (it is not the case that whenever there is a long pause it coincides with a syntactic boundary). The fact that Dutch 6-month-old infants do not rely on pitch cues so heavily suggests that this reliance is not merely the result of weighting still being driven by a holistic sensitivity (Johnson & Seidl, 2005, in press). On the contrary, by this age infants must have started to pay attention to the distribution of cues in their native language, but it is the developmental trajectory towards weighting of local cues that allows Dutch infants to rely on pause. However, it is unlikely that 6-month-old infants have reached the end of the continuum from global to local cues. Indeed, Seidl (2007) shows that 6-month-old infants, unlike adults, are not able to exploit a single cue to segment the speech stream, but need the convergence of either pause or preboundary lengthening with pitch in order to succeed at the task. This finding seems to indicate that although

6-month-olds use a more local processing strategy than 4-month-old infants, they are still not able to exploit local cues as much as adults (Aasland & Baum, 2003).

Another important difference that arises from the comparison between 4- and 6-month-olds is that in the present experiment we found a significant listening preference for the non-clause in the Baseline condition. This is not what was found for 6-month-old infants in Nazzi, Kemler Nelson *et al.* (2000) or in the replication of that experiment in Seidl (2007). In both of these studies, 6-month-old infants showed a listening preference for the clausal unit in the test phase of the experiment. However, it may be that younger infants respond differently to clausal units than older ones. The literature on familiarity and novelty preferences (e.g. Hunter & Ames, 1988) has documented changes in preference due to length of exposure and complexity of the stimuli. Since the length of familiarization was the same across the two ages, the reason for the change of preference can only lie in the other factor.[4] The present stimuli could be more complex for 6-month-olds than 4-month-olds because the older infants are beginning to engage in other tasks, namely defining the phonological inventory of their language and segmenting words (Bortfeld *et al.*, 2005; Werker & Tees, 1984). Another related reason why these stimuli would be more complex at the older age is that 4-month-olds attend only to global characteristics, not needing to pay attention to each cue independently.[5] Nevertheless, the fact that young infants preferred one of the versions to the other in this clause segmentation task suggests that they are already exploiting prosodic cues for this task, a prerequisite for the prosodic bootstrapping of clauses.

It is not surprising that 4-month-olds act in a different manner from 6-month-olds. Apart from the longer perceptual exposure to the ambient language, 4- and 6-month-olds also differ in their language production abilities which may further impact their speech perception. Specifically, infants at these two ages differ in their use of prosodic cues in their vocalizations. For example, evidence suggests that the use of pitch within utterances varies as a function of age. Analysis of the acoustic properties of 3-month-olds' utterances reveals that they typically involve a falling pitch (Kent & Murray, 1982), a tendency that has been explained on the basis of physiology by Lieberman (1967). Similarly, Nathani, Oller and Cobo-Lewis (2003) suggest that 3-month-old

infants exploit the cue of preboundary lengthening because it is biologically determined and hence present in infants' vocalizations regardless of whether infants receive any auditory input; that is, both hearing and deaf 3-month-old infants produce preboundary syllables that are significantly longer than non-final syllables. In contrast with these biologically driven prosodic productions in young infants, the intonational pattern used by infants over 5 months of age in reduplicative babble reflects that of the ambient language; for instance, French-learning infants use more rising contours than English-learning ones (Whalen, Levitt & Wang, 1991). In other words, the use of pitch as marker of boundaries seems to be primarily driven by physiological pressures in young infants, while older infants seem to have already started to tune into the intonational patterns of their input language.

Furthermore, this developmental change is in line with other language discrimination studies suggesting that infants become attuned to the rhythmic characteristics of their language by 5 months of age (Nazzi, Jusczyk & Johnson, 2000). Newborns have been shown to be able to discriminate languages with different rhythmic properties, regardless of whether one of them is their native language (Nazzi, Bertoncini & Mehler, 1998); however, they are unable to discriminate two languages from the same rhythmic class (Ramus, 2002) and the same is true for 2–3-month-old infants (Christophe & Morton, 1998). On the other hand, 4- and 5-month-olds are able to discriminate two languages belonging to the same rhythmic class provided one of them is their native language (Bosch & Sebastián-Gallés, 1997; Nazzi, Jusczyk & Johnson, 2000).

In conclusion, this study has contributed to elucidate how infants are able to begin to find clauses in their native language. Our results suggest that 4-month-old infants are applying a global or holistic strategy that is not sensitive to differences in cue-weighting in the ambient language. This strategy allows them to find clausal boundaries in running speech, which constitutes the scaffold on which infants are able to construct a cue-weighting scheme that is appropriate to the ambient language at least by the time they are 6 months. This hypothesis predicts that infants will rely on all three cues regardless of ambient language at 4 months of age, but that they would converge, at least partially, on a cue-weighting appropriate to their language at a later age. Further research with other linguistic populations and at other ages may provide support for this prediction.

[4] Alternatively, it could be that 4-month-old infants are less familiar with grammatical units in the input language than older infants. For this reason they may show a sensitivity to the more prosodically distinct edges (more distinct in the sense of having larger changes in duration, pitch and pause within the word sequence) of grammatical units, which would eventually allow them to track how reliable each prosodic cue to clausal boundaries is in their ambient language.

[5] As suggested by one reviewer, the null results in the manipulation conditions may be due to the fact that absence of a cue made the task harder for the infants and this increased difficulty caused infants to show responses which hover at the border between these two kinds of preference.

References

Aasland, W., & Baum, S. (2003). Temporal parameters as cues to phrasal boundaries: a comparison of processing by left-hemisphere-damaged and right-hemisphere-damaged individuals. *Brain and Language*, **87**, 385–399.

Bahrick, L.E. (2002). Generalization of learning in three-and-a-half-month-old infants on the basis of amodal relations. *Child Development*, **73**, 667–681.

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

Beckman, M., & Hirschberg, J. (1994). *The ToBI annotation conventions*. Ohio State University. (Manuscript with examples)

Beckman, M., & Pierrehumbert, J.B. (1986). Japanese prosodic phrasing and intonation synthesis. In *Proceedings of the 24th ACL* (pp. 173–180). New York.

Boersma, P., & Weenik, D. (2005). Praat: Doing phonetics by computer (version 4.0.13) [Computer program]. Retrieved 26 May 2005, from //www.praat.org/.

Bortfeld, H., Morgan, J.L., Golinkoff, R.M., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science*, **16** (4), 298–304.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in four-month-old infants from monolingual and bilingual environments. *Cognition*, **65**, 33–69.

Bruce, G., & Gårding, E. (1978). A prosodic typology for Swedish dialects. In E. Gårding, G. Bruce, & R. Bannert (Eds.), *Nordic prosody* (pp. 219–228). Lund: Lund University.

Cashon, C., & Cohen, L. (2004). Beyond u-shaped development in infants' processing of faces: an information-processing account. *Journal of Cognition and Development*, **5** (1), 59–80.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, **1**, 215–219.

Colombo, J., Mitchell, D.W., Coldren, J., & Freeseman, L. (1991). Individual differences in infant visual attention: are short lookers faster processors or feature processors? *Child Development*, **62**, 1247–1257.

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

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, **4**, 597–650.

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, **266**, 718–719.

Freeseman, L., Colombo, J., & Coldren, J. (1993). Individual differences in infant visual attention: discrimination and generalization of local and global properties at 4 months of age. *Child Development*, **64**, 1191–1203.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, **15**, 103–113.

Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access: II. infant data. *Journal of Memory and Language*, **51**, 547–567.

Hayes, B. (1990). Precompiled phrasal phonology. In S. Inkelas & D. Zec (Eds.), *The phonology–syntax connection* (pp. 85–108). Chicago, IL: University of Chicago Press.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, **28**, 377–396.

Hernandez-Reif, M., & Bahrick, L.E. (2001). The development of visual-tactual perception of objects: amodal relations provide the basis of learning arbitrary relations. *Infancy*, **2** (1), 51–72.

Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P.W., Cassidy, K.W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, **26**, 269–286.

Hirst, D., & Di Cristo, A. (1998). *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.

Horne, M., Strangert, E., & Heldner, M. (1995). Prosodic boundary strength in Swedish: final lengthening and silent interval duration. In K. Elenius & P. Branderud (Eds.), *Proceedings of the International Congress of Phonetic Sciences* (pp. 170–173). Stockholm.

House, D., Hermes, D., & Beaugendre, F. (1998). Perception of tonal rises and falls for accentuation and phrasing in Swedish. In R.H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the International Congress of Phonetic Sciences* (pp. 2799–2802). Sydney.

Houston, D., & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, **26** (5), 1570–1582.

Hunter, M.A., & Ames, E.W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, **5**, 69–95.

Johnson, E., & Seidl, A. (2005). Cross-linguistic differences in 6-month-olds' clause segmentation strategies. Poster presented at ISCA Workshop on Plasticity in Speech Perception, London.

Johnson, E., & Seidl, A. (in press). A cross-linguistic perspective on infants' detection of junctures in speech. *Infancy*, **13** (5).

Jusczyk, P.W. (1989). Perception of cues to clausal units in native and non-native languages. Paper presented at the biennial meeting of the Society of Research in Child Development, Kansas City.

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

Jusczyk, P.W. (2003). American infants' perception of cues to grammatical units in non-native languages and music: evidence from Polish and Japanese. In D. Houston, A. Seidl, G. Hollich, E. Johnson, & A.M. Jusczyk (Eds.), *Jusczyk lab final report*. Retrieved from http://hincapie.psych.purdue.edu/Jusczyk: The Johns Hopkins University.

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

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

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

Kent, R.D., & Murray, A.D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, **72**, 353–365.

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

Lieberman, P. (1967). *Intonation, perception and language*. Cambridge, MA: MIT Press.

Mayo, C., & Turk, A. (2004). Adult–child differences in acoustic cue weighting are influenced by segmental context: children are not always perceptually biased toward transitions. *Journal of the Acoustical Society of America*, **115**, 3184–3194.

Morgan, J., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Erlbaum.

Nathani, S., Oller, D.K., & Cobo-Lewis, A.B. (2003). Final syllable lengthening (FSL) in infant vocalizations. *Journal of Child Language*, **30**, 3–25.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology*: *Human Perception and Performance*, **24** (3), 1–11.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, **21**, 779–784.

Nazzi, T., Jusczyk, P.W., & Johnson, E. (2000). Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. *Journal of Memory and Language*, **43**, 1–19.

Nazzi, T., Kemler Nelson, D., 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.

Nittrouer, S., Manning, C., & Meyer, G. (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *Journal of the Acoustical Society of America*, **94**, S.1865.2.

Nittrouer, S., & Miller, M.E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, **101**, 2253–2266.

Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press/Bradford Books.

Pinker, S. (1991). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Quinn, P., Burke, S., & Rush, A. (1993). Part–whole perception in early infancy: evidence for perceptual grouping produced by lightness similarity. *Infant Behavior and Development*, **16**, 19–42.

Quinn, P., & Johnson, M. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, **1**, 31–46.

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

Sanderman, A., & Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech*, **40**, 391–409.

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. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, **57**, 24–48.

Seidl, A., & Johnson, E. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, **9** (6), 565–573.

Seidl, A., Johnson, E., Redman, A., & Brentari, D. (2004). Segmentation of clauses in English, Dutch, and American sign language. Paper presented at the annual Boston University Conference on Language Development, Boston, MA.

Smiljanic, R. (2006). Early vs. late focus: pitch-peak alignment in two dialects of Serbian and Croatian. In L. Goldstein, D.H. Whalen, & C.T. Best (Eds.), *Laboratory phonology 8* (pp. 495–518). Berlin: Mouton de Gruyter.

Soderstrom, M., Seidl, A., Kemler Nelson, D., & Jusczyk, P.W. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language*, **49**, 249–267.

Vaissiere, J. (1983). Language-independent prosodic features. In A. Cutler & D. Ladd (Eds.), *Prosody: Models and measurements* (pp. 53–66). Berlin: Springer-Verlag.

Venditti, J. (1995). *Japanese ToBI labeling guidelines*. Ohio State University. (Manuscript with examples)

Volskaya, N. (2003). Virtual and real pauses at clause and sentence boundaries. In *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions.

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

Whalen, D.H., Levitt, A.G., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, **18**, 501–516.

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

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.