

REPORT

Infant word segmentation revisited: edge alignment facilitates target extraction

Amanda Seidl¹ and Elizabeth K. Johnson²

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

2. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Abstract

In a landmark study, Jusczyk and Aslin (1995) demonstrated that English-learning infants are able to segment words from continuous speech at 7.5 months of age. In the current study, we explored the possibility that infants segment words from the edges of utterances more readily than the middle of utterances. The same procedure was used as in Jusczyk and Aslin (1995); however, our stimuli were controlled for target word location and infants were given a shorter familiarization time to avoid ceiling effects. Infants were familiarized to one word that always occurred at the edge of an utterance (sentence-initial position for half of the infants and sentence-final position for the other half) and one word that always occurred in sentence-medial position. Our results demonstrate that infants segment words from the edges of an utterance more readily than from the middle of an utterance. In addition, infants segment words from utterance-final position just as readily as they segment words from utterance-initial position. Possible explanations for these results, as well as their implications for current models of the development of word segmentation, are discussed.

Introduction

In order to interpret utterances an infant first must identify the words they contain. As any non-native listener can attest, this is not a trivial task, especially when we consider the variability of word forms, speakers, and dialects and that, even in infant-directed speech (IDS), words are not reliably presented in isolation (van de Weijer, 1998; Woodward & Aslin, 1990). To make matters more complex, infants often have only brief contextual exposure to a new word, hearing it perhaps only once or twice in a given conversational period.

Despite these obstacles, there is evidence that young infants can segment words from fluent speech (Jusczyk & Aslin, 1995). After familiarization with two passages containing multiple tokens of a single target word such as *cup* or *dog*, infants as young as 7.5 months listened longer to isolated tokens of these words than to structurally similar words. More specifically, infants who were familiarized to passages containing *cup* and *dog* listened longer to lists of the words *cup* and *dog* than to lists of the unfamiliar words *bike* and *feet*. Infants who were familiarized to passages containing *bike* and *feet* showed the opposite listening preference. Each familiarization

passage included the target word uttered twice in utterance-initial position, twice in utterance-medial position and twice in utterance-final position.

Infants' precocious ability to segment words from speech has resulted in a large number of segmentation studies designed to understand how infants parse fluent speech into words. Certainly phonotactic cues are helpful in word segmentation for adults and for infants 9 months and older (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993; Mattys & Jusczyk, 1994), but younger infants know little about the phonotactic constraints of their input language and thus must rely on other cues present in the signal. Segmentation cues that seem to be available to younger infants include lexical stress pattern (Curtin, Mintz & Christiansen, 2005; Johnson & Jusczyk, 2001; Jusczyk, Houston & Newsome, 1999), statistical/co-occurrence relations among syllables (Saffran, Aslin & Newport, 1996), and allophonic/subphonemic cues to word boundaries (Johnson, 2003; Johnson & Jusczyk, 2001; Mattys & Jusczyk, 2001). In this paper, we focus on an additional segmentation strategy: We investigate whether utterance boundaries also play an important role in infants' early segmentation attempts.

Address for correspondence: Amanda Seidl, Department of Speech, Language and Hearing Sciences, Purdue University, W. Lafayette, IN 47906, USA; e-mail: aseidl@purdue.edu

The Edge Hypothesis

Research has demonstrated that given adequate familiarization, 7.5- to 10-month-olds readily segment words from utterance-medial position (Johnson, 2005; Jusczyk *et al.*, 1999; Mattys & Jusczyk, 2001; Saffran *et al.*, 1996). However, these findings do not rule out the possibility that words located at the beginning or end of an utterance are easier to extract from running speech than words falling in the middle of an utterance. Indeed, there is little evidence to speak to this possibility because the great majority of infant segmentation studies follow Jusczyk and Aslin's (1995) original design in which familiarization passages contain words located utterance-finally or -initially as well as utterance-medially. In these studies, it is impossible to determine if infants segment words from utterance-final position more readily than they segment words from utterance-medial position (see, for example: Bortfeld, Morgan, Golinkoff & Rathbun, 2005; Houston & Jusczyk, 2000; Johnson, 2003; Polka & Sundara, 2003). In this paper, we propose the *Edge Hypothesis*. This hypothesis states that infants use utterance edges to help them locate words in fluent speech. If the Edge Hypothesis is correct, we predict that infants should segment words from utterance edges more readily than they segment words from utterance-medial position.

Evidence motivating the Edge Hypothesis

There is ample evidence that in both infant- and adult-directed speech (ADS), words falling along utterance boundaries are acoustically more salient than words embedded in the middle of an utterance (Fougeron & Keating, 1997; Keating, Cho, Fougeron & Hsu, 2003; Nespor, Guasti & Christophe, 1996; Quene, 1992, 1993; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992). Moreover, these effects seem to be more prominent in IDS than ADS (Bernstein Ratner, 1986; Broen, 1972; Cruttenden, 1986; Fernald & Mazzie, 1991; Fisher & Tokura, 1996). Furthermore, given the preponderance of short utterances in IDS, utterance edges are simply more frequent in IDS than in ADS (Phillips, 1973). Taken together, these findings suggest that IDS may be well designed to facilitate the use of an utterance boundary segmentation strategy (for a useful discussion of segmentation and IDS see Kemler Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989; Thiessen, Hill & Saffran, 2005).

Additional motivation for the Edge Hypothesis comes from adult speakers' behavior. It has been shown that American, as well as Turkish, mothers tend to place new words that they are teaching their infants at the end of

sentences, even if this means producing ungrammatical sentences (Aslin, Woodward, LaMendola & Bever, 1996). Aslin *et al.* (1996) suggested that general processing constraints such as recency may constrain the processing of auditory input and focus the child on the edges of utterances. Attention to the ends of linguistic units is also predicted by Slobin's acquisition principle to 'pay attention to the ends of things' (Slobin, 1973). In a Connectionist model of early word segmentation, Aslin *et al.* (1996) reported better segmentation of words that occurred utterance-finally. Aslin *et al.* (1996) argue that once an infant identifies a word+utterance boundary, he/she can use this to find utterance-internal word boundaries. Easier segmentation at utterance edges is also predicted by the INCDROP model of word segmentation (Dahan & Brent, 1999). Proponents of the INCDROP model argue that edge effects are due to computations in the INCDROP framework, but they readily admit that such edge effects could also be due to perceptual saliency effects or general processing constraints.

Finally, Aslin (1999) provided behavioral evidence suggesting that infants are attuned to the ends of utterances by testing 8-month-olds' segmentation of nonsense utterances. Aslin's results suggested that infants segment nonsense items more readily when they occurred in utterance-final as opposed to utterance-medial position. Similarly, Golinkoff, Hirsh-Pasek and Alioto (1995) found that 21- to 36-month-olds learn new words better when those words are in utterance-final position than utterance-medial position and recently, Cummings and Fernald (2003) reported on similar findings. Still other studies indicate that listeners' attention to the ends of units persists into adulthood (Conway & Christiansen, 2005; Golinkoff & Alioto, 1995).

The initial vs. the final edge

The Edge Hypothesis, as we have presented it, is location-free: It makes no prediction as to whether the initial or final edge of an utterance provides a more useful segmentation cue. However, there may be good reason to predict that one edge is more useful than the other. On the one hand, parental input (Aslin *et al.*, 1996), as well as behavioral studies (Aslin, 1999), suggest that utterance-final position is the privileged location as far as ease of segmentation is concerned. Work on phrasal prominence also supports the notion that utterance-final position should be particularly salient in English, since English with its SVO word order tends to place its sentential accent on the last word in an utterance (Cinque, 1993; Nespor *et al.*, 1996). Although more work has focused on the final edge, findings in the adult word recognition literature might predict that the initial edge is a more useful

segmentation cue than the final edge. For example, word onsets seem to play an important role in lexical access, which is most likely related to the fact that speech unfolds in an initial to final manner (Marslen-Wilson, 1990; Spivey-Knowlton, Sedivy, Eberhard & Tanenhaus, 1994; Swingley, 1997; Swingley, Pinto & Fernald, 1999). In addition, Newport, Gleitman and Gleitman (1977) provided evidence that children may pay particular attention to the beginnings of sentences and that the particular nature of mothers' IDS at the beginnings of sentences is correlated with their children's language growth. There is also some evidence that second language learners may find it easier to segment words from sentence-initial position than from sentence-final position (Barcroft & VanPatten, 1997). Thus, although there is good reason to suspect that utterance boundaries might help infants segment words from speech, there is no clear evidence as to whether the initial or final edge is more privileged in terms of ease of segmentation. Investigating the ease of segmentation at both edges is an important endeavor both within and across languages given that different word classes occur in different positions for different languages. For example, in English transitive sentences, noun objects tend to occur in utterance-final position and in Japanese transitive sentences, verbs tend to occur in utterance-final position.

Testing the Edge Hypothesis

In this experiment we used the same Headturn Preference Procedure as Jusczyk and Aslin (1995) to test whether 7.5- to 8.5-month-old infants segment words that are located along utterance edges more readily than they segment words that are located utterance-medially. As a secondary question, we asked whether words falling along the initial and final edge of an utterance are equally easy to extract from speech. One passage contained a target CVC word that always occurred at an utterance edge (the initial edge for infants in Condition A and the final edge for infants in Condition B) and the other passage contained a target CVC word that always occurred utterance-medially. Familiarization was limited to 25 seconds for each familiarization passage rather than the usual 45 seconds in order to avoid ceiling effects, and maximize the odds that we would see a performance difference between the utterance-edge and utterance-medial words. If infants segment words that fall along utterance boundaries more readily than they segment utterance-medial words, then infants should have greater orientation time difference scores (mean orientation time to familiar test items minus mean orientation time to unfamiliar test items) for edge words than medial words.

Method

Participants

Sixty-four normally hearing English-learning 8-month-olds (average age was 7 months and 30 days with a range of 7 months and 12 days to 8 months and 15 days; 33 males, 31 females) from the Lafayette, Indiana region were tested. An additional 14 infants were excluded for the following reasons: 10 were excluded for fussing or crying and four were excluded because of experimenter errors.

Stimuli

Four monosyllabic CVC words (two words and two nonsense words) were chosen for use as target stimuli: *geff* [gɛf], *deeb* [dib], *cash* [cæʃ], and *pig* [pig]. Each of the four words was recorded in an utterance-initial paragraph, an utterance-medial paragraph, and an utterance-final paragraph (a sample of the exact sentences appears in the Appendix). In this experiment all utterances corresponded not only to a single Intonational Phrase, but also to a sentence/single clause. Thus, we cannot address whether syntactically defined units differ in any relevant way from prosodically defined units in this study. Each of these paragraphs was six sentences long, and read in an infant-directed style by a native English speaker from Arkansas who was naïve to the purpose of the experiment. Passages were matched for length at an average of 20 s. Fifteen isolated tokens of each of the four words were subsequently recorded for use during the test phase (see Table 1 for acoustic measurements of target words in all four positions). These word lists were matched at an average of 17.5 s. All recordings were made in a sound-shielded booth using CSL software and hardware. The sentences were digitized

Table 1 *Measurements of words*

Target word	Position	Vowel duration	Initial segment duration	f0	Amplitude
cash	initial	354 ms	61 ms	725.5 Hz	68 db
cash	medial	191 ms	73 ms	552 Hz	71 db
cash	final	397 ms	63 ms	432 Hz	69 db
cash	citation	354 ms	75 ms	358 Hz	70 db
deeb	initial	151 ms	24 ms	510 Hz	62 db
deeb	medial	168 ms	29 ms	425 Hz	65 db
deeb	final	347 ms	37 ms	392 Hz	65 db
deeb	citation	495 ms	35 ms	382 Hz	70 db
geff	initial	131 ms	37 ms	667.5 Hz	72 db
geff	medial	141 ms	25 ms	459 Hz	75 db
geff	final	386 ms	30 ms	322 Hz	69 db
geff	citation	386 ms	42 ms	386 Hz	70 db
pig	initial	101 ms	32 ms	603 Hz	67 db
pig	medial	123 ms	73 ms	507 Hz	73 db
pig	final	454 ms	44 ms	428 Hz	72 db
pig	citation	444 ms	76 ms	357 Hz	70 db

at a 22 KHz sampling rate. Digitized speech samples were transferred to a Macintosh G4 computer for playback during the experiment. The average loudness of the speech samples as measured by a sound meter was 71 dB.

The average pauses between utterances were measured and shown to be .486 seconds long. We also measured the pause in our utterance-types to see if there was a longer pause when the target appeared at an edge than when it did not. When the target word was utterance-initial, the average pause duration was .418 seconds. In addition, we measured the duration of the pause (or lack of pause) before the target word in medial and final utterances. Here we found that in medial utterances the pause duration was an average of .088 seconds before the target word and in final utterances the pause duration was an average of .081 seconds before the target word. Thus, clearly, if pause is a key cue to aid in segmentation, infants would have an easier time with words at the beginning of utterances.

It is important to mention that our utterance-medial target words did occur at syntactic boundaries; in fact, the lexical Noun Phrase itself (e.g. *geff*) was a boundary. In addition most of the target words were within larger lexical projections and occurred at either the initial or final edge of these projections; 14 out of 24 of the target words were within a Verb Phrase, five were within an Adjective Phrase and the remaining five were within functionally headed phrases. Given that lexically headed phrases tend to map well to prosodic units (Seidl, 2000) (namely Phonological Phrases in the sense of Selkirk, 1984) then we can assume that the words at the edges of these syntactic units were also at the edges of prosodically perceptible units.

Design

Out of the 64 participants, 32 of the infants received Condition A familiarization (one word in utterance-initial position and one word in utterance-medial position) and the other 32 received Condition B familiarization (one word in utterance-final position and one word in utterance-medial position; see Table 2 for details). Further, half of the infants were familiarized with *cash* and *deeb* and the other half were familiarized with *geff* and *pig*. All infants were tested on the same four test items: *cash*, *deeb*, *pig*, and *geff*.

All infants were familiarized to two passages until they reached an orientation time criterion of 25 seconds. Once the infant accumulated this target amount of orientation time the test phase began. During the test

phase, all infants heard three blocks of trials. There were a total of four trials per block. Two of these trials contained lists of unfamiliar stimuli and two contained lists of familiar stimuli (those target words that were presented in the familiarization passages). A different random ordering was used for the test trials in each block.

Apparatus

A computer controlled the presentation of the speech samples and recorded an observer's coding of the infants' looking preferences via 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 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 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 & Aslin, 1995; Kemler-Nelson *et al.*, 1995) 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 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 side light and specific auditory stimulus were made randomly via computer program. Both the experimenter and the caregiver wore tight-fitting headphones (Peltor Aviation

Table 2 *Design of experiment*

	Group 1	Group 2	Group 3	Group 4
Condition A	cash initial, deeb medial	deeb initial, cash medial	geff initial, pig medial	pig initial, geff medial
Condition B	cash final, deeb medial	deeb final, cash medial	geff final, pig medial	pig final, geff medial

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 central 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 (chosen randomly by the computer program used to run the experiment) 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 oriented 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 oriented away was not counted in the overall tally.

The experiment consisted of two phases: Experiment and test. During familiarization, two passages played alternately until the infant accrued 25 seconds of orientation time to each passage. During the test phase infants heard word lists that contained the target words presented in the familiarization passages or novel words not presented in the familiarization passages. Thus, if an infant in Condition A, Group 1 heard the words *cash* and *deeb* in familiarization, then the unfamiliar words lists for these infants were lists containing *geff* and *pig*. 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 familiar test items and the other two contained unfamiliar test items. The dependent measure was the average orientation time across trials to each stimulus type.

Results

In order to ascertain whether familiarization time could affect infants' preference for medial- versus edge-aligned words, we ran a paired *t*-test comparing infants' accumulated orientation time to passages containing utterance-aligned target words ($M = 13.56$; $SD = 4.4$) versus passages containing utterance-medial target words ($M = 13.83$; $SD = 3.6$). This test revealed that infants listened to these two types of familiarization passages equally long, $t(63) = .49$, $p > .1$. Thus, our results could not be attributed to greater attention to passages containing edge-aligned words than passages containing utterance-medial words. We also checked to ensure that infants did not hear many more tokens of the utterance-initial targets than the utterance-medial and utterance-final targets. Tabulations

of the exact number of times infants heard words of each type during the familiarization revealed that infants were exposed to similar numbers of all word types: For infants in the final-edge condition, utterance-aligned words were heard an average of 11.5 times ($SD = 3.6$) and medial words were heard an average of 13 times ($SD = 2.7$). A *t*-test revealed no significant difference between the number of times infants heard final-edge utterance aligned words and medial words, $t(31) = 1.9$, $p > .05$. For infants in the initial-edge condition, utterance aligned words were heard an average of 13 times ($SD = 2.4$) and medial words were heard an average of 14 times ($SD = 2.9$). A *t*-test revealed no significant difference between the number of times infants heard initial-edge aligned words and medial words, $t(31) = 1.7$, $p > .05$. Finally, we found no correlation ($r(64) = .88$, $p > .1$) between the number of tokens of initial and final word heard during the familiarization and orientation times to familiar versus unfamiliar words.

In order to examine performance in the test phase of the experiment, mean orientation time to Familiar Edge test items ($M = 10.02$; $SD = 3.4$), Familiar Medial test items ($M = 8.67$; $SD = 3.4$), and Unfamiliar test items ($M = 8.75$; $SD = 3.10$) were calculated for all 64 infants (see Figure 1).

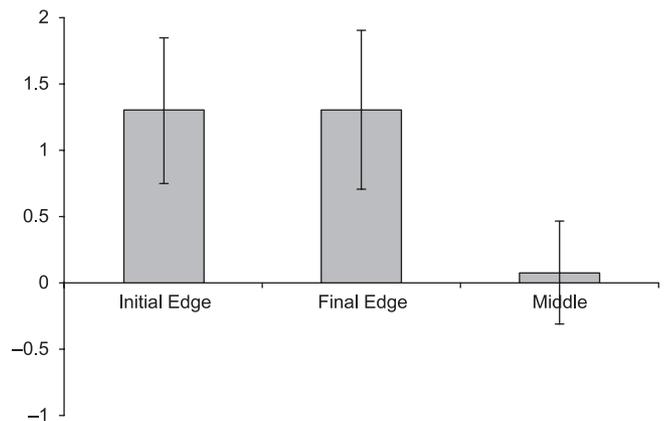


Figure 1 Mean difference scores, calculated by subtracting the mean orientation time to unfamiliar test items from the mean orientation time to familiar test items, are indicated above. More specifically, the mean orientation time difference for the initial domain edge was calculated by subtracting the mean orientation time to unfamiliar test items by those infants familiarized with utterance-initial words from those same infants' mean orientation time to familiar utterance-initial words. The mean orientation time difference to final domain edge items was calculated in a similar manner. The mean orientation time difference to medial words includes orientation time to familiar medial minus unfamiliar test items for all infants, i.e. for those infants familiarized with utterance-initial words as well as for those infants familiarized with utterance-final words. Positive mean difference scores represent longer orientation to familiar test items during the test phase.

Table 3 Means and SDs for initial and final edges

	Edge familiar	Middle familiar	Unfamiliar
Initial	10 s (SD = 3.8)	8.0 s (SD = 3.3)	8.7 s (SD = 3.3)
Final	10.1 s (SD = 2.9)	9.3 s (SD = 3.4)	8.8 s (SD = 3.0)

(The orientation time means and SDs for Initial and Final Edge test items are shown in Table 3.)

A mixed Design ANOVA, 3 (Test Item Type: Familiar Edge, Familiar Medial, Unfamiliar) \times 2 (Condition: Final Edge and Initial Edge), revealed a significant main effect of Test Item Type, $F(2, 62) = 6.2$, $p < .005$. There was no significant main effect of Condition, $F(1, 62) < 1$, or significant interaction between Test Item Type and Condition, $F(2, 62) < 1$. Planned comparisons revealed that infants oriented significantly longer to Familiar Edge test items than Unfamiliar test items, $t(63) = 3.11$, $p < .01$. Infants also oriented significantly longer to Familiar Edge test items than Familiar Medial test items, $t(63) = 2.74$, $p < .01$. In contrast, infants oriented equally long to Familiar Medial test items and Unfamiliar test items, $t(63) < 1$. Infants also oriented equally long to familiar Initial edge words and familiar Final edge words, $t(62) < 1$. These results suggest that infants segment words from utterance edges more readily than from utterance middles. A similar, yet much less theoretically interesting explanation for our results is that the infants recognized the words they had heard at the utterance edges more readily than they recognized words that had occurred utterance-medially because the utterance-edge targets were a closer acoustic match to the isolated test items than the utterance-medial targets. We find this explanation unlikely for two reasons: (1) past studies have shown that infants easily recognize words in isolation even if they have never heard the word in isolation before (and vice versa) and (2) the utterance-initial targets were often much more acoustically distinct from the isolated test items in terms of vowel duration, VOT, and pitch height, than utterance-final targets, and yet infants performed identically on utterance-initial and utterance-final items, although, as pointed out by one reviewer, it is possible that these results would not hold up had we chosen to use a longer familiarization time.

Discussion

Our results suggest that given brief exposure to new words, infants segment words located along the edges of utterances more readily than they segment words located in the middle of utterances. These findings support earlier work suggesting that utterance-final words are

optimally positioned to facilitate segmentation (Aslin *et al.*, 1996; Aslin, 1999). In addition, we found that positioning a target word at the onset of an utterance facilitates segmentation just as much as positioning a target word at the end of an utterance. In the future, it will be fruitful to examine whether (and at what age) the facilitatory nature of edge alignment extends to utterance-internal junctures such as phrase boundaries (see Christophe, Mehler & Sebastián-Gallés, 2001, and Christophe, Gout, Peperkamp & Morgan, 2003, for a discussion of infants' sensitivity to utterance-internal boundaries) and clause boundaries (Nazzi, Kemler Nelson, Jusczyk & Jusczyk, 2000; Soderstrom, Seidl, Kemler Nelson & Jusczyk, 2003).

Another important finding of this study is that locating target words in utterance-initial or utterance-final position seems to have equally facilitatory effects on ease of segmentation. In short, our findings support a location-free version of the Edge Hypothesis. This finding is somewhat surprising given past reports that utterance-final position should be the easiest location for word segmentation. Although it is possible that a ceiling effect may explain our failure to find any evidence that utterance-final words are easier to segment from speech than utterance-initial words, the fact remains that both positions seem to be highly privileged with respect to ease of segmentation. Only future experimentation will enable us to determine if in fact an even shorter familiarization time might reveal that final position facilitates segmentation better than initial position.

An additional issue warranting further investigation is whether our results would hold up for common mass nouns such as *bread* or even nominal forms of verbs such as *running*. Although this study cannot directly address this issue, we find it highly unlikely that the outcome of these experiments would have differed had we used mass nouns rather than names given that these two word classes are not prosodically distinct and infants at this age do not yet possess the syntactic or semantic competence to distinguish them. More questionable is how similar results would be in another language with larger differences between common and proper names. For example, in Swahili where proper names are not prefixed, but common nouns have nominal class prefixes, both can occur in initial and final position, but the additional morphology may make segmentation of the prefixed words a more complex task. One additional cross-linguistic implication of this study related to word class, is that infants exposed to languages with looser word order (e.g. Chinese) and more variable word classes at utterance edges may have early vocabularies that reflect a wider variety of word classes (as suggested in the recent work of Tardif, Shatz & Naigles, 1997), than infants

exposed to languages which have stricter word order and less variable early vocabularies (e.g. English).

Returning to the issue of the possible effect of word class on segmentation in this study, it is also important to note that while much of the literature on word segmentation has focused on common as opposed to proper nouns (Höhle & Weissenborn, 2003), the little work that has been done on infants' segmentation of words other than nouns suggests that word class may play an important role in how easily words can be segmented from speech. For example, English-learning infants' ability to segment verbs from speech is delayed with respect to the ability to segment nouns (Nazzi, Dille, Jusczyk, Shattuck-Hufnagel & Jusczyk, 2005). Our results also highlight the importance of considering word order in language acquisition. Interestingly, in English, both utterance-initial and utterance-final position may be occupied by proper nouns and function morphemes. These types of lexical units have been proposed to play an important role in early language acquisition (Bortfeld *et al.*, 2005; Shi, Werker & Cutler, 2006; Shi, Cutler, Werker & Cruickshank, 2006). Prior work has suggested that names and function morphemes are highly frequent items in the child's input, and could thus help infants break speech stream into linguistically relevant units. The results of the current study suggest that in addition to their frequency of occurrence, positioning may also help to make these lexical units readily accessible to the language-learning infant. And recognition of these highly frequent lexical items in fluent speech could help infants break up the speech stream into smaller units (Brent, 1999).

As discussed in the introduction, there are different types of explanations for why infants would find it easier to segment words at utterance edges, none of which are mutually exclusive. First, infants could have a general cognitive bias to attend to the edges of units. This bias could possibly be related to the recency and primacy effects seen in free recall studies. Another possibility is that words at utterance boundaries are easier to segment because they are perceptually salient due to prosodic modifications such as initial strengthening and final lengthening. Yet another possibility is that the long silence preceding or following the target word is the primary reason why infants segment words at utterance boundaries more readily than they segment words from the middle of utterances. Regardless of why infants segment utterances the way they do, the fact remains that utterance edges appear to be a hot spot for segmentation activity.

One interesting implication of this finding is that the usefulness of utterance boundaries could be interpreted as indirect evidence for the facilitatory role of IDS in early word segmentation. Infant-directed speech consists of

mostly short multi-word utterances produced in a slow rate with exaggerated pitch excursions (Bernstein Ratner, 1986; Broen, 1972; Cruttenden, 1986; Fernald & Mazzie, 1991). Recent research has demonstrated that the exaggerated pitch contours of IDS most likely facilitates word segmentation (Thiessen *et al.*, 2005). The present study presents another dimension along which IDS could facilitate segmentation. The fact that utterances are shorter in IDS than ADS indicates that words are more likely to be flanked by an utterance boundary in IDS than in ADS. Thus, since utterance boundaries facilitate segmentation, the high frequency of utterance boundaries in IDS may make it easier for infants to extract words from IDS than ADS.

In summary, this study has revealed that both the final and the initial edge of utterances are privileged with respect to ease of segmentation. Additional work will be needed to more fully understand the implications of the Edge Hypothesis. The Edge Hypothesis could be strengthened by further experimentation and isolation of the specific speech cues involved at utterance boundaries and an exploration of the extent to which infants rely upon one cue or another. In addition, it will be important to explore the degree to which infants younger than 7.5 months may segment words from speech when they are aligned with utterance boundaries. If segmentation is found at a younger age with words aligned with the utterance edge, this could advance our understanding of how the infant's word segmentation skills begin to develop. In the meantime, these findings shed light on how infants begin to tackle the word segmentation problem by showing that initially infants segment words from the edges of utterances more readily than they segment words from the middles of utterances.

Appendix

Sample passages with utterance-initial target words

Geff runs the big circus in Toronto. Geff wants to juggle out there all day long. Geff has a great time making sugar snaps. Geff would love to learn to bake lasagna. Geff could do a great job fixing our school. Geff hires clowns to drive a tiny car around.

Sample passages with utterance-final target words

At the circus I like low geff. It is such bad luck to have a rough geff. We'll have to tell her that we know geff. The lion tamer knows how to ensure geff. The elephants always want to thank geff. The clowns start dancing when they say geff.

Sample passages with utterance-medial target words

I like how geff runs the circus. I wonder if geff wants to juggle too. We know geff has a great time. We're sure geff would love to learn. I think geff could do a great job. They say geff hires clowns all year.

Acknowledgements

We would like to thank Ilana Feld, Amanda Klein and Lindsey Leidig for help in organizing and recruiting participants. In addition, we are grateful to Lisa Goffman, George Hollich and two anonymous reviewers for much useful advice and are equally thankful to the parents and babies who participated in these studies. All mistakes are, of course, our own. This work is supported by NICHD R03 HD046463-01.

References

- Aslin, R. (1999). Utterance-final bias in word recognition by 8-month-olds. Poster presented at the Society for Research in Child Development, Albuquerque.
- Aslin, R., Woodward, J., LaMendola, N., & Bever, T. (1996). Models of word segmentation in fluent maternal speech to infants. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134). Mahwah, NJ: Lawrence Erlbaum Associates.
- Barcroft, J., & VanPatten, B. (1997). Acoustic salience of grammatical forms: the effect of location, stress, and boundedness on Spanish L2 input processing. In W.R. Glass & A.T. Perez-Leroux (Eds.), *Contemporary perspectives on the acquisition of Spanish, vol. 2: Production, processing, and comprehension* (pp. 109–121). Somerville, MA: Cascadilla.
- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics*, **14**, 303–309.
- 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.
- Brent, M. (1999). Speech segmentation and word discovery: a computational perspective. *Trends in Cognitive Sciences*, **3**, 294–300.
- Broen, P. (1972). The verbal environment of the language-learning child. *ASHA Monographs*, **17**.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: the role of prosody. *Journal of Phonetics*, **31**, 585–598.
- Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Infants' sensitivity for prosodic boundary cues. *Infancy*, **2**, 385–394.
- Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry*, **24**, 239–297.
- Conway, C., & Christiansen, M. (2005). Modality-constrained statistical learning of tactile, visual and auditory sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **31** (1), 24–39.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Cummings, A., & Fernald, A. (2003). Is it easier for infants to learn a new word first presented in isolation or in a multiword utterance? Paper presented at the biennial meeting of the Society of Research in Child Development, Tampa, Florida.
- Curtin, S., Mintz, T., & Christiansen, M. (2005). Stress changes the representational landscape: evidence from word segmentation. *Cognition*, **96**, 233–262.
- Dahan, D., & Brent, M. (1999). On the discovery of novel word-like units from utterances: an artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology*, **128**, 165–185.
- Fernald, A., & Mazzei, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, **27**, 209–221.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development*, **67**, 3192–3218.
- Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, **101**, 3728–3740.
- Golinkoff, R., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: implications for language acquisition. *Journal of Child Language*, **22** (3), 703–726.
- Golinkoff, R., Hirsh-Pasek, K., & Alioto, A. (1995). Lexical items are learned better in infant-directed than adult-directed speech. Paper presented at the Boston Conference on Language Development.
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, **6**, 122–127.
- Houston, D., & Jusczyk, P. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, **26** (5), 1570–1582.
- Johnson, E.K. (2003). Word segmentation during infancy: the role of subphonemic cues to word boundaries. Unpublished doctoral dissertation, The Johns Hopkins University.
- Johnson, E.K. (2005). English-learning infants' representations of word-forms with iambic stress. *Infancy*, **7** (1), 95–105.
- Johnson, E.K., & Jusczyk, P. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language*, **44**, 548–567.
- Jusczyk, P., & Aslin, R. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1–23.
- Jusczyk, P., Friederici, A., Wessels, J., Svenkerud, V., & Jusczyk, A.M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, **32**, 402–420.
- Jusczyk, P., Houston, D., & Newsome, M. (1999). The beginnings

- of word segmentation in English-learning infants. *Cognitive Psychology*, **39**, 159–207.
- Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in laboratory phonology VI* (pp. 143–161). Cambridge: Cambridge University Press.
- Kemler Nelson, D., Hirsh-Pasek, K., Jusczyk, P., & Cassidy, K.W. (1989). How prosodic cues in motherese might assist language learning. *Journal of Child Language*, **16**, 55–68.
- Kemler Nelson, D.G., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A., & Gerken, L. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development*, **18**, 111–116.
- Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.
- Mattys, S., & Jusczyk, P. (1994). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, **78**, 91–121.
- Mattys, S., & Jusczyk, P. (2001). Do infants segment words or continuous recurring patterns? *Journal of Experimental Psychology: Human Perception and Performance*, **27**, 644–655.
- Nazzi, T., Dilley, L., Jusczyk, A.M., Shattuck-Hufnagel, S., & Jusczyk, P. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, **48** (3), 279–298.
- Nazzi, T., Kemler Nelson, D., Jusczyk, P., & Jusczyk, A.M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. *Infancy*, **1**, 123–147.
- Nespor, M., Guasti, M., & Christophe, A. (1996). Selecting word order: the rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin: Akademie Verlag.
- Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother i'd rather do it myself: some effects and non-effects of maternal speech style. In C.E. Snow & C.A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–150). Cambridge: Cambridge University Press.
- Phillips, J. (1973). Syntax and vocabulary of mothers' speech to young children: age and sex comparisons. *Child Development*, **44**, 182–185.
- Polka, L., & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. In M.J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona.
- Quene, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics*, **20** (3), 331–350.
- Quene, H. (1993). Segment durations and accent as cues to word segmentation in Dutch. *Journal of the Acoustical Society of America*, **94** (4), 2027–2035.
- Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.
- Seidl, A. (2000). *Minimal indirect reference: A theory of the syntax-phonology interface*. New York: Routledge.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Shi, R., Werker, J.F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, **10**, 187–198.
- Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, **119** (6), EL61–EL67.
- Slobin, D.I. (1973). Cognitive prerequisites from the development of grammar. In C.A. Ferguson & D.I. Slobin (eds.), *Studies of child language development*. New York: Holt, Rinehart and Winston.
- Soderstrom, M., Seidl, A., Kemler Nelson, D., & Jusczyk, P. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language*, **49**, 249–267.
- Spivey-Knowlton, M., Sedivy, J., Eberhard, K., & Tanenhaus, M. (1994). Psycholinguistic study of the interaction between language and vision. In *AAAI-94 Workshop on Integration of Natural Language and Vision Processing* (pp. 189–192). Seattle, WA: American Association for Artificial Intelligence.
- Swingle, D., Pinto, J.P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, **71**, 73–108.
- Swingle, D. (1997). Word recognition and representation in young children. Unpublished doctoral dissertation, Stanford University.
- Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *Journal of Child Language*, **24**, 535–565.
- Thiessen, E., Hill, E., & Saffran, J. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, **7** (1), 53–71.
- van de Weijer, J. (1998). *Language input for word discovery*. Max Planck Series in Psycholinguistics 9.
- 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.
- Woodward, J., & Aslin, R. (1990). Segmentation cues in maternal speech to infants. *International Conference on Infant Studies*, Montreal.

Received: 18 August 2005

Accepted: 25 January 2006