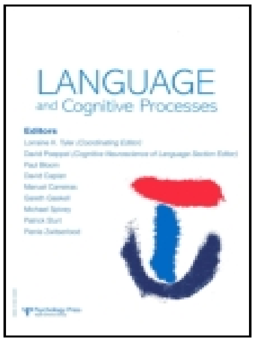

Language and Cognitive Processes

ISSN: 0169-0965 (Print) 1464-0732 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp20

Colliding cues in word segmentation: The role of cue strength and general cognitive processes

Daniel J. Weiss , Chip Gerfen & Aaron D. Mitchel

To cite this article: Daniel J. Weiss , Chip Gerfen & Aaron D. Mitchel (2010) Colliding cues in word segmentation: The role of cue strength and general cognitive processes, Language and Cognitive Processes, 25:3, 402-422, DOI: 10.1080/01690960903212254

To link to this article: http://dx.doi.org/10.1080/01690960903212254

Published online: 28 Sep 2009.

LANGUAGE AND COGNITIVE PROCESSES
2010, 25 (3), 402–422

Colliding cues in word segmentation: The role of cue strength and general cognitive processes

Daniel J. Weiss
Department of Psychology and Program in Linguistics, Pennsylvania State University, University Park, PA, USA

Chip Gerfen
Department of Spanish, Italian, and Portuguese and Program in Linguistics, Pennsylvania State University, University Park, PA, USA

Aaron D. Mitchel
Department of Psychology, Pennsylvania State University, University Park, PA, USA

The process of word segmentation is flexible, with many strategies potentially available to learners. This experiment explores how segmentation cues interact, and whether successful resolution of cue competition is related to general executive functioning. Participants listened to artificial speech streams that contained both statistical and pause-defined cues to word boundaries. When these cues 'collide' (indicating different locations for word boundaries), cue strength appears to dictate the predominant parsing strategy. When cues are relatively equal in strength, the ability to successfully deploy a segmentation strategy significantly correlates with stronger performance on the Simon task, a non-linguistic cognitive task typically thought to involve executive processes such as inhibitory control and selective attention. These results suggest that general information processing strategies may play a role in solving one of the early challenges for language learners.

Keywords: Language acquisition; Speech segmentation; Statistical learning; Simon task.

Correspondence should be addressed to Daniel Weiss, 643 Moore Building, University Park, PA 16802. E-mail: djw21@psu.edu

http://www.psypress.com/lcp DOI: 10.1080/01690960903212254

INTRODUCTION

The task of speech segmentation consists of two fundamental challenges. Learners must identify cues in the speech stream that signal word boundaries, and they must weight these cues according to their effectiveness, a process which results in the development of language specific segmentation strategies (e.g., Cutler, Mehler, Norris, & Segui, 1986). Despite the fact that first language learners appear to acquire segmentation abilities naturally and with ease, the task itself is more complex than one might intuit. A central challenge lies in the fact that languages themselves lack invariant cross-linguistic cues to guide learners to correctly segment the input stream (Klatt, 1979). In simple terms, the lack of invariant cues to segmentation presents learners with an extremely large set of potential cues, while at the same time they lack *a priori* knowledge of which cue, or constellation of cues, may be the most reliable for successful segmentation in any given language. This situation approximates a 'combinatorial explosion' (von der Malsburg, 1995) of potential computations, a large problem space from which the learner needs to converge on a correct learning strategy. A growing body of recent research has established a role for statistical learning, a domain-general mechanism (e.g., Hunt & Aslin, 2001; Saffran, Johnson, Aslin, & Newport, 1999), in providing a solution to this longstanding problem (Newport & Aslin, 2004; Saffran, Aslin, & Newport, 1996a; Saffran, Newport, & Aslin, 1996b; Thiessen & Saffran, 2003). The logic of this approach is that statistical learning provides a foothold into the segmentation dilemma by allowing the learner to use statistical cues to boundaries as a means of bootstrapping to other language-specific segmentation cues (e.g., Saffran, 2003). Once the learner has shifted to using multiple segmentation cues (as suggested by Morgan & Saffran, 1995, for example), the learner's task then becomes weighting the various language particular cues, attending to those that are more effective while discounting less effective cues. The consensus, in fact, suggests that the process of speech segmentation is integrative, as learners consolidate information across many available cues (e.g., Brent & Cartwright, 1996; Christiansen, Conway, & Curtin, 2005; Jusczyk, Houston, & Newsome, 1999; Mattys, Jusczyk, Luce, & Morgan, 1999; Morgan & Saffran, 1995; Myers, Jusczyk, Kemler Nelson, Charles-Luce, Woodward, & Hirsh-Pasek, 1996).

Given the potentially foundational role that a non-linguistic, general cognitive ability such as statistical learning may play in the initial developmental stages of speech segmentation, the question explored in this paper is whether other domain general cognitive processes may also enhance our understanding of how humans segment speech. Specifically executive function studies of selective attention and inhibitory control have investigated how people attend to relevant cues while ignoring irrelevant cues (e.g., Downing, 2000; Hasher, Zacks, & May, 1999). If cue integration, as we note above, requires both the ability to selectively focus on particular cues while inhibiting or discounting other acoustic information, it stands to reason that this type of processing may be modulated by individual differences in executive function. Here we begin to address this question by examining whether performance on a colliding cue segmentation task (a linguistic task in which several segmentation cues are available but indicate conflicting word boundaries) correlates with performance on the Simon task, a test of general executive functioning traditionally described as measuring inhibitory control or selective attention (Hedge & Marsh, 1975; Simon & Berbaum, 1990; see Lu & Proctor, 1995 for review). If there is a relationship between domain general cognitive processing reflected in the Simon task and linguistic processing reflected in the segmentation task, then under appropriate experimental manipulations, we predict that individual differences in executive function will correlate with differential performance on the segmentation task.

Colliding cues in segmentation

In order to track the relative values of particular segmentation cues, it is useful to evaluate their contribution in isolation. At the same time, given the integrative nature of speech segmentation, it is difficult to directly assess the contributions of individual cues. One empirical technique employed to address this issue is the 'colliding cue' methodology (e.g., Johnson & Jusczyk, 2001; Mattys, 2004; Mattys et al., 1999; Thiessen & Saffran, 2003). Here, two unaligned cues are set into opposition, each indicating word boundaries at a different location in the speech input. Using this method to study the segmentation problem during language acquisition, for example, Johnson and Jusczyk (2001) investigated whether infants prefer to segment according to 'speech cues' (such as stress and coarticulation) or transitional probability cues (of the sort used in Saffran et al., 1996a, 1996b). Johnson and Jusczyk (2001) found that at 8 months of age, infants prefer to segment the stream using the stress and coarticulation cues, thus concluding that these cues are 'ranked' higher than the transitional probability cues.

Interestingly, a follow-up study by Thiessen and Saffran (2003) showed that this pattern failed to hold for 7-month-old infants, who instead demonstrated a preference for using transitional probability cues to segment an artificial speech stream, even when stress cues were available (and dictated word boundaries at a separate location). Thiessen and Saffran (2003) speculated that the reversal may be due to the infants using transitional probability cues as a means of discovering language-particular stress patterns, at which point they switch to a strategy in which stress is weighted more heavily. Once the infants discover that stress cues 'are not infallible', the authors speculate they may again alter their segmentation strategy (at around 11 months of age; see discussion in Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003). Taken together, the results from these experiments demonstrate that during the course of development, segmentation strategies are flexible and may evolve as new information about the speech stream is gathered.

Importantly, as infants switch their segmentation strategies, their ability to deploy the non-preferred strategy does not disappear. This is evidenced, for example, by adults' continued ability to use transitional probabilities to segment speech streams (e.g., Saffran et al., 1996b), and experimental evidence suggests that infants and adults exhibit similar constraints in performing statistical computations (Newport & Aslin, 2004). Indeed, even the flexibility demonstrated by infants while learning to segment also persists into adulthood after segmentation has been mastered. Studies manipulating the background noise level during adult segmentation tasks (Mattys, 2004; Mattys, White, & Melhorn, 2005) clearly show that adult learners are capable of deploying different segmentation strategies as environmental affordances change. Such results indicate that the hierarchical weighting of different segmentation cues and strategies is a feature of both language learners and mature language users (see Mattys et al., 2005). This feature allows learners to deal with noise emanating from the environment, as well as with the noise intrinsic to selecting a segmentation strategy from a set of partially overlapping cues. Failure to resolve the noise from competing cues may, in fact, result in a lack of learning, as demonstrated by Toro-Soto, Rodríguez-Fornells, and Sebastián-Gallés (2007).

In sum, the emerging picture from the colliding cue literature suggests that the process of learning to segment speech relies on integration across multiple cues. This process requires learners to selectively tune to and attend to particularly relevant cues, and discount information from less informative potential segmentation cues. It follows, then, that individual differences in executive function relating to selective attention and inhibitory control might be expected to correlate with performance on a linguistic colliding cue task, such that learners who are better able to inhibit potentially contradictory cues should perform better in cue conflict conditions than learners who are less able to exercise inhibitory control. With this prediction in mind, we presented participants in our experiment with a colliding cue segmentation task and subsequently tested each with a visual version of the Simon task, as described below.

The Simon task

The Simon task is a non-linguistic cognitive task that measures the effect of conflicting cues on information processing (Simon & Small, 1969). In the Simon task, participants must attend to a particular relevant dimension of the stimulus (e.g., colour) while other irrelevant dimensions are also present (e.g., spatial position) that either support or oppose the attended cue. The task typically consists of having participants make speeded forced-choice button press responses while viewing a stimulus on the screen which is either on the same side as the appropriate response button (in congruent trials) or on the opposite side of the appropriate response button (in incongruent trials). Despite the irrelevance of location information in the task, many studies have demonstrated and replicated a reaction time advantage for congruent relative to incongruent trials, a result known as the Simon effect. Performance on the Simon task is thought to reflect selective attention, inhibitory control, or response switching costs (Hedge & Marsh, 1975; Simon & Berbaum, 1990; see Lu & Proctor, 1995 for review).

Experimental logic

To test the hypothesis that the general cognitive processes such as selective attention or inhibitory control are related to linguistic processes involved in resolving colliding cues, we presented learners with a colliding cue speech segmentation task followed by a Simon task. The colliding cue task required learners to segment a speech stream that contained two opposing segmentation cues. Consistent with previous studies of segmentation in which learners segment artificial languages, we employed transitional probabilities as one of our segmentation cues. Our second cue was prosodic, consisting of pauses. As alluded to above, pauses do not systematically serve to mark word junctures in speech. However, their presence in the speech stream is known to provide an extremely salient cue for the learner (e.g., Fisher & Tokura, 1996; Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss, & Kennedy, 1987; Mattys & Clark, 2002). In fact, recent work on statistical learning has focused on the contribution of pauses to the triggering of rule learning mechanisms, the suggestion being that when pauses are congruent with distributional cues, learners are able to generalise beyond the surface input structures to extract underlying rule-based patterns (Mueller, Bahlmann, & Friederici, 2008; Peña, Bonatti, Nespor, & Mehler, 2002). Pauses as brief as 25 ms have been shown to elicit such changes in behaviour (Peña et al., 2002; cf. Perruchet, Tyler, Galland, & Peereman, 2004), again underscoring the fact that pauses in the speech stream play a central role in early language acquisition, *including* the process of word segmentation. In addition, from the perspective of the experimenter, pauses are also easy to manipulate with extreme precision. Here, we varied pause durations across different conditions and predicted, based on work from other domains (e.g., Massaro, 1998), that reliance on a particular cue would increase as a function of cue strength. Previous efforts have explored cue strength by manipulating environmental influences on segmentation (see Mattys, 2004, Mattys et al., 2005). Thiessen and Saffran (2004) also explored cue strength in a study that used a similar logic to our experimental design. They presented a colliding cue task with transitional probabilities and spectral tilt cues, a secondary correlate of one of the three known cues that comprise stress, at different locations in an artificial speech stream. They found that 9-month-old infants treated the spectral tilt cues differently from 1-year-old and adult learners. However, we note that their experimental method was different from our study in two important ways. First, as the authors acknowledge, the cue manipulated in the Thiessen and Saffran study is not known to be a naturally occurring cue to stress. Second, there was no gradient manipulation of the stress cue. In this experiment, we present the first effort to directly manipulate the strength of the cues themselves in order to ascertain how the strength of particular cues affects the outcome of cue collisions in a statistical learning paradigm (see Mattys & Melhorn, 2007 for a demonstration of cue strength effects for segmentation of known words in sentence context). If the segmentation process is indeed based on graded contributions from many potential cues, then it is reasonable to predict that as the strength of a particular cue increases, learners will be more likely to parse the stream using that cue. This point has never been directly tested using the colliding cue paradigm implemented in developmental studies of word segmentation. Consequently, the results of such a study can both inform and refine our understanding of how segmentation cues interact. For example, when language learners encounter a collision between stress and transitional probabilities, what factors contribute to the weighting of the cues? Given that learners ultimately develop language-appropriate segmentation strategies that differ across languages (Cutler et al., 1986), the logical conclusion is that individual cues will be assigned different weights based on their respective 'strength' (i.e., prominence) within a given language.

As mentioned above, this colliding cue task was followed by a Simon task. Importantly, the Simon task may share a subset of underlying cognitive abilities that are useful for successfully deploying a segmentation strategy in a colliding cue paradigm. As noted above, the Simon task requires learners to attend to a colour cue (e.g., blue or red) while ignoring positional information (e.g., left or right). This may be achieved by either selective attention or inhibitory control (Lu & Proctor, 1995). In the early processes of speech segmentation, learners must identify viable segmentation cues (such as transitional probability cues) while ignoring other cues that are not reliable indicators of word boundaries (such as short duration pauses). Furthermore, the process of cue integration itself may involve processes such as selective attention and inhibitory control in order to adjust the weighting of individual cues.

If a relationship exists between the cognitive demands required resolve colliding cues in segmentation and the Simon task, then performance on the segmentation task should correlate with performance on the Simon task in conditions where pauses and transitional probability cues are highly competitive. This prediction is restricted to highly competitive conditions because greater competition has been shown to require more top-down processing (e.g., Torralbo & Beck, 2008) and is therefore more likely to engage selective attention and/or inhibitory control. Such a relationship between our tasks would be consistent with interactive theories of language acquisition and general cognitive learning mechanisms (e.g., Cairns, Shillcock, Chater, & Levy, 1997) and may provide insight into the well-known fact that the course of language acquisition varies considerably across individuals (Bates, Bretherton, & Snyder, 1991; see General Discussion).

METHOD

In a between-subjects design with three experimental conditions, we inserted pauses of differing lengths within statistically defined words, yielding an artificial speech stream in which pauses and statistical information in the form of transitional probabilities dictate word boundaries at distinct locations. Our control condition consisted of a speech stream with only transitional probability cues. Following the segmentation task, we had participants perform a visual version of the Simon task to determine whether a measure of general executive functioning correlated with successful segmentation in the competing cue situation.

Participants

One hundred and sixty-one experimentally naïve undergraduates participated. All were monolingual English, Introductory Psychology students (98 female and 63 male) participating for class credit. Four additional participants were excluded from the analysis (two who failed to follow instructions, and two due to technical failure).

Stimuli

Participants heard an artificial language, shown to be learnable in previous experiments (Weiss, Gerfen, & Mitchel, 2009), consisting of four trisyllabic words (consonant-vowel (CV) pairs: CV.CV.CV) re-synthesised from natural speech. One male voice was digitally recorded while producing CVC syllables (e.g., [bab], etc.). The CV sequences were recorded with coda consonants in order create labial, alveolar, palatoalveolar, and velar VC transitions. CV syllables were also recorded with no coda consonant in order to provide natural sounding word and foil word endings for the trisyllabic strings presented in isolation in the testing phase. Tokens were hand-edited in Praat

(Boersma, 2001) to control for duration, normalised in SoundForge© to control for loudness, and resynthesised with an identical f0 contour on each CV syllable to control for pitch differences across syllables.

The four words were concatenated pseudo-randomly into a continuous stream, with each presented an equal number of times and with no word following itself. To preserve statistical integrity, each word was followed by every other word an equal number of times. There were no additional acoustic cues to word boundaries in the stream.

In the control condition, the speech stream was presented without pauses between syllables. In all experimental conditions, pauses were inserted between the second and third syllables of the statistically defined trisyllabic words (see Figure 1). Thus, statistical and pause cues to segmentation collided, with each indicating different word boundaries. Three experimental conditions varied with respect to the inserted pause length: Condition 1 – 25 ms, Condition 2 – 50 ms, and Condition 3 – 75 ms. Participants were randomly assigned to one of the conditions.

The stimuli controlled for both syllable-to-syllable transitional probabilities and segment-to-segment transitions (defined as consonant-to-vowel-to-consonant, etc.; Newport, Weiss, Aslin, & Wonacott, 2004). Figure 1 shows the words comprising the language. Each had perfect (1.0) within-word syllable-to-syllable transitional probabilities and 0.33 probabilities at word boundaries. Since no words were repeated, any word could be followed by one of three other words. Segment-to-segment statistics were also consistent; within-word transitions were 0.5 and dipped to 0.33 at word boundaries.

Procedure

Participants received verbal instructions to listen to an audio stream and were informed that they would be tested on information learned from the stream. The stream was presented through headphones on computers with iTunes™ software and consisted of three 4-minute blocks with a 1-minute silence between blocks. Participants were monitored by the experimenters from an adjacent room.

After the listening task, participants completed a 16-item, two-alternative forced-choice test. Word pairs were presented with 1 second of silence

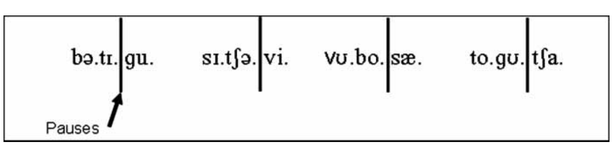

Figure 1. Placement of the pauses in the experimental conditions relative to the statistically defined words. Note that the pause cues and statistical cues each indicate a different location for potential word boundaries.

between choices, and with 4 seconds of silence between each pair. Each statistically defined word was tested four times, presented with two different part-words (counterbalanced for order), for a total of 16 trials. The part-words were created by concatenating the last syllable of one word with the first two syllables of a second word (as is traditionally used in these types of experiments). In our experimental conditions, the statistically defined part-words corresponded to pause-defined words, since the pauses were inserted in between the second and third syllables of each statistically defined word. Participants circled either '1' or '2' to indicate which string they believed to be a word from the stream. Participants also completed a questionnaire assessing language background (assessing bilingualism, number of languages spoken, years spoken) as well as self-ratings on effort and confusion.

After the segmentation task, participants performed a visual version of the Simon task. The visual stimuli were presented on PCs using E-Prime (Psychology Software Tools, 2002). In the task, participants sat in front of a screen and were instructed to focus on a fixation point (a small cross measuring approximately 1° by 1°) located in the centre of the screen. During test trials, the fixation point disappeared and a square (measuring approximately 1° by 1°) of either red or blue colour appeared, and remained on screen for 1500 ms if there was no response.[1] Participants were instructed to press a response key that corresponded to the box's colour. The blue key was located on the left side of the keyboard and the red key was located on the right side. The position of the square varied by trial, with the square appearing on the left side, the right side, or in the central position of the screen. Thus, the task-relevant information was colour (since that dictated which button should be pressed). The task-irrelevant feature was position (regardless of where the square appeared on the screen, it was the colour of the square that dictated the button-press). During test, the task-relevant colour information was either congruent with the task-irrelevant information of position (e.g., a blue square appears on the left side of the screen, and the blue left button had to be pressed) or incongruent (e.g., a blue square appears on the right side of the screen, and the blue left button had to be pressed). Participants received on-screen feedback following their response or following the expiration of the time limit. Subjects began their test session with 24 practice trials followed by 126 test trials.

Scoring and analysis

In each segmentation test-trial, participants chose between two trisyllabic items. In the control condition, the choice was between a statistical word and

[1] Limiting the time window for response is typical in Simon tasks (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004).

a statistical part-word. Due to the insertion of the pause between the second and third syllable of statistically defined words (see Figure 1), in the experimental conditions, although the test was identical to the control condition, the choices now represented statistically defined words versus pause-defined words. As a consequence, it is important to note that in this experiment, deviations from chance in either direction indicated learning. For consistency with previous studies, in this paper above-chance performance indicated that participants followed statistical cues in segmenting the stream. Alternatively, scoring below chance indicated that participants followed the pause cues.

Due to this property of the experimental data (that deviations away from chance in either direction indicated learning), we correlated the Simon test scores with the *absolute value* of each test score away from chance in the cue collision conditions (in this case chance was calculated as 8, since the test was a 16-item two-alternative forced-choice task). Thus, our correlation indicated whether the participant's ability to successfully deploy a segmentation strategy (either using pauses *or* transitional probability cues) was related to performance on the Simon task. As a secondary analysis, we also compared the slopes across the experimental conditions. Though not ideal as a primary analysis, the slope, as a less sensitive measure to high variability, can provide a complement to the validity of the correlational measure.

For the Simon task, we recorded whether responses were accurate as well as reaction time measurements. As is typical for scoring the Simon task (see Grosjean & Mordkoff, 2002), neither incorrect trials nor recovery trials (i.e., trials that immediately followed an incorrect response) were included in the analysis. Additionally, we discarded correct trials that were further than two standard deviations from the mean as outliers. The Simon effect was then calculated by subtracting the reaction times of congruent trials from those of the incongruent trials.

RESULTS

Overall, the average number of statistically defined words learned varied by condition (i.e., pause duration). The mean score (out of 16) in the baseline condition (no pauses) was 10.18 (2.68); in Condition 1 (25 ms pauses) it was 8.79 (2.68);[2] in Condition 2 (50 ms pauses) it was 7.98 (3.99); and for Condition 3 (75 ms pauses) it was 3.77 (3.32; see Table 1). In the baseline condition, learning was significantly above chance, $t(39) = 5.14$, $p < .001$. In

[2] We initially conducted two separate types of tests (across subjects), one with pauses inserted in the statistically defined words during test, and another omitting them. We found no significant differences between test types and subsequently pooled the data in our results.

TABLE 1
Experimental results: Mean score on the segmentation task (out of 16).

	No pause (baseline)	*25 ms pause*	*50 ms pause*	*75 ms pause*
Mean	10.18	8.79	7.98	3.77
SD	2.68	2.68	3.99	3.32
N	40	43	43	35

Condition 1, learning approached significance (i.e., parsing according to statistically defined words), $t(42) = 1.94$, $p = .060$. In Condition 2, learning was at chance levels, $t(42) = -0.04$, $p = .970$ and in Condition 3, learning was significantly below chance (i.e., learners successfully segmented the language following pause-defined words), $t(34) = -7.51$, $p < .001$. An ANOVA showed a significant main effect for pause length, $F(3, 157) = 27.01$, $p < .001$ (see Figure 2). A post-hoc Bonferroni comparison revealed a significant difference in group means between the 75 ms condition and all other conditions ($p < .001$, for all comparisons), as well as a difference between the baseline condition and the 50 ms condition ($p = .013$). There was

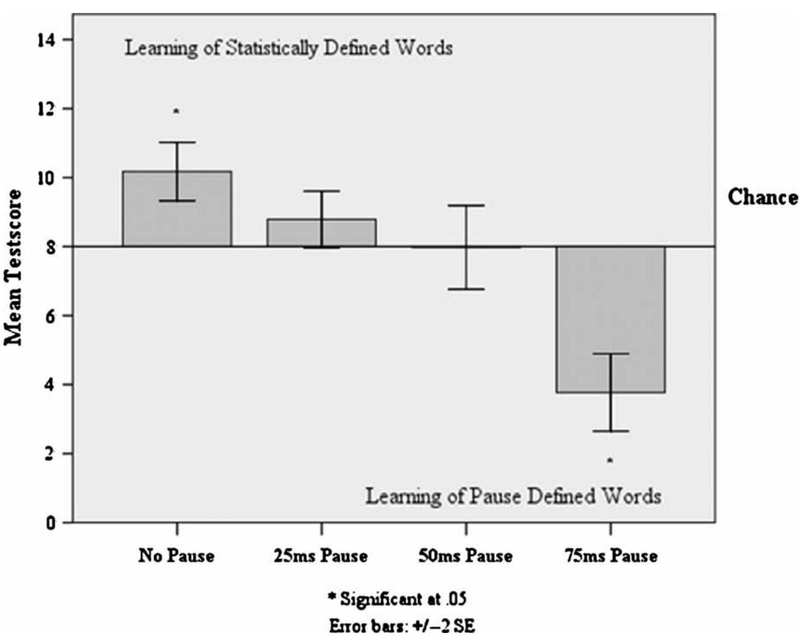

Figure 2. Results of the segmentation task. The y-axis plots the scores from a 16-item two-alternative forced choice task. Excepting the control condition, scoring above chance indicates that learners parsed according to statistics, whereas below-chance scores indicate learners parsed according to the pauses (see text).

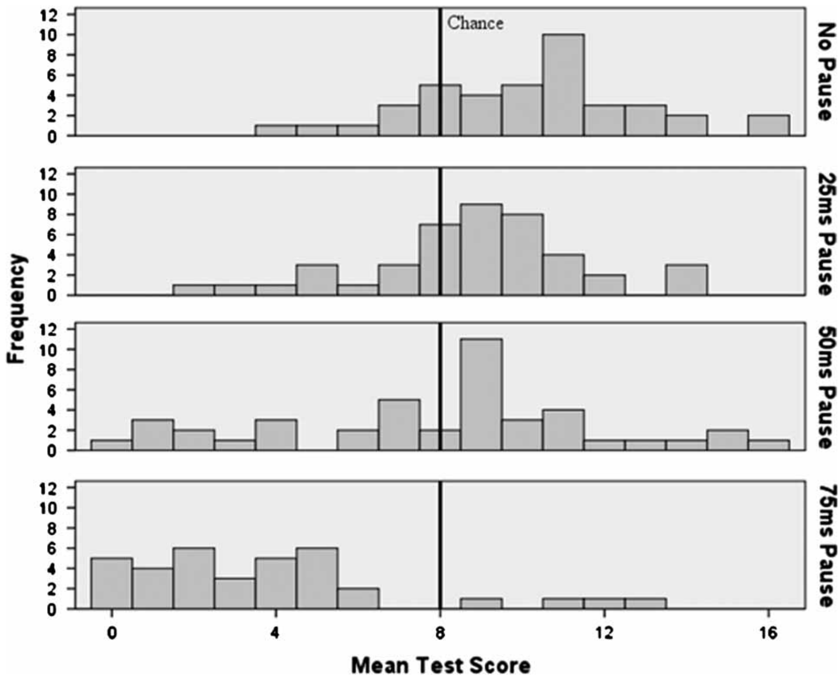

Figure 3. Histograms of the distribution of scores on the 16-item test for all four conditions.

a significant negative linear trend in mean test score by pause length, $F(1, 158) = 65.28$, $p < .001$.

Figure 3 illustrates the distribution of responses across conditions. In Condition 1 the majority of participants parsed according to statistically defined words, while in Condition 3 the majority of participants parsed according to pause-defined words. In Condition 2, although many participants performed at chance levels, there were also many learners who scored at either of the extremes, successfully segmenting the language by closely following pause-defined words or statistically defined words.

The mean Simon effect[3] by condition was: baseline, 36.44 ms ($SD = 19.34$); Condition 1, 37.61 ms ($SD = 21.00$); Condition 2, 38.21 ms ($SD = 19.69$); Condition 3, 35.47 ms ($SD = 21.58$). A one-way ANOVA revealed no significant difference in mean Simon effect between conditions, $F(3, 134) = 0.13$, $p = .944$. There were thus no significant differences in performance across conditions.

The performance data from the segmentation task across all conditions (using the *absolute value* away from chance, see above) were compared with

[3] A typical mean Simon effect is around 30 ms (see Grosjean & Mordkoff, 2002).

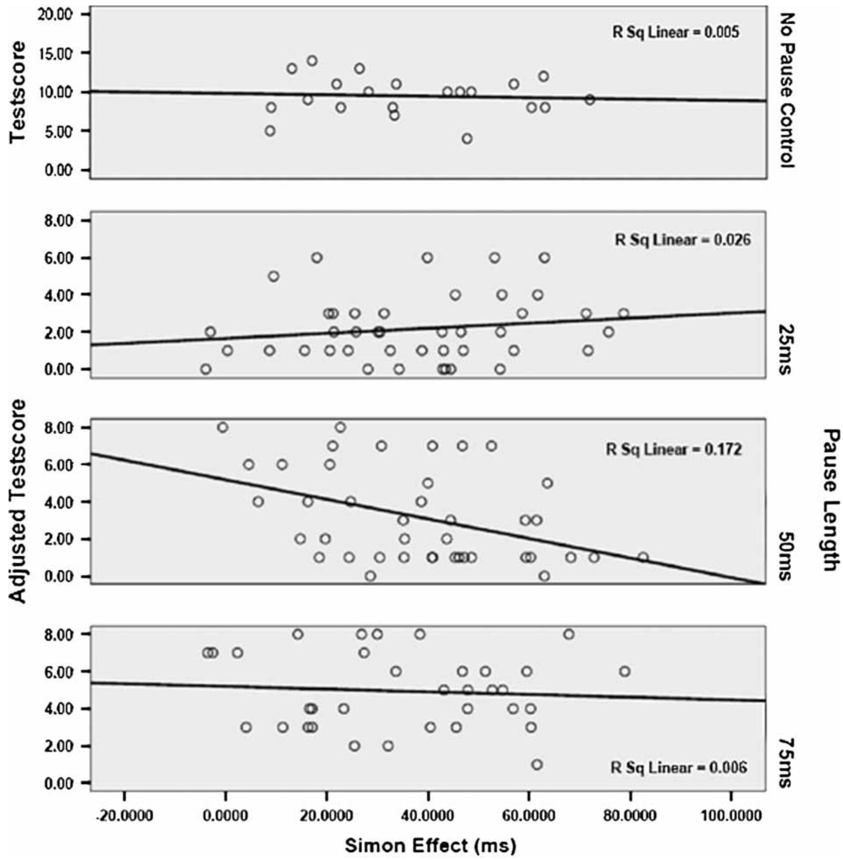

Figure 4. Scatterplots showing the relationship between participants' performance on the Simon and segmentation tasks.

the results of the Simon task. A low Simon effect indicates less interference from the incongruent positional information, whereas a high Simon effect indicates more interference. In the baseline condition[4] as well as Conditions 1 and 3, we found no significant correlation between performance on the segmentation and Simon tasks, $r(19) = -.070$, $p = .761$, $B = -.009$; $r(40) = .160$, $p = .310$, $B = .014$; $r(32) = -.077$, $p = .664$, $B = -.007$, respectively;[5] see Figure 4). However, Condition 2 exhibited a significant negative correlation between learning and the Simon effect, $r(39) = -.415$, $p = .007$,

[4] In the baseline condition, performance was compared with the unadjusted testscore, as performance was unidirectional. In the other three conditions, the Simon scores were compared with an adjusted test score (the absolute difference of the test score from chance).

[5] For the first 17 baseline participants, we did not conduct the Simon test.

$B = -.053$. Specifically, in Condition 2 participants with better inhibitory control were more likely to achieve a consistent and thus successful parsing strategy (by either following the statistically defined or the pause-defined words), whereas learners performing worse on the Simon task were more likely to perform at chance on the segmentation task. Across all three experimental conditions, the correlation between the Simon task and performance on the segmentation task was not significant, $r(117) = -.141$, $p = .129$.

In order to compare the slopes of the three regression lines, we conducted a homogeneity of regression analysis across the three experimental conditions that revealed a marginally significant difference among the three slopes, $F(2, 114) = 2.81$, $p = .064$. A comparison of individual slopes between conditions found that Condition 1 and Condition 2 were significantly different, $t(79) = 3.02$, $p = .003$, while a comparison of Condition 2 and Condition 3 similarly yielded a strong trend toward significance, $t(71) = 1.91$, $p = .060$. In contrast, there was no significant difference in slopes between Condition 1 and Condition 3, $t(72) = 1.02$, $p = .312$.

GENERAL DISCUSSION

Two main conclusions can be drawn from the data obtained here. We found that changing the pause length across conditions yielded a graded change in performance (see Figure 2). As pause length increased, learners were progressively more likely to segment according to pause boundaries. This indicates that the strength of a segmentation cue can directly impact the learner's choice of segmentation strategy. Most notably, when segmentation cues were relatively equally matched in strength (as evidenced by scores on the segmentation task), learners who were able to segment the stream successfully (using either available cue) tended to score better on the Simon task, whereas those at chance tended to score worse on the Simon task. Because the Simon task is a general cognitive task that measures the effect of conflicting cues on information processing (see Lu & Proctor, 1995), this result suggests a relationship (discussed below) between the fundamental mechanisms underlying language acquisition and general information processing mechanisms.

The correlation in performance between the Simon task and the colliding cue segmentation task that emerged in Condition 2 supports an association between the early mechanisms of language acquisition (statistical learning of word boundaries) and general executive functioning involved in resolving conflicting cues. This finding is interesting in light of the similarities and differences across tasks. The colliding cue word segmentation task is an auditory, linguistic task that requires learners to process distributional information (either in the form of pauses or transitional probabilities) over time. The Simon task is a visual task that does not require learners to maintain information beyond the immediate trial. Another difference between tasks is that in the Simon task, the required response and the irrelevant stimuli share a dimension, namely location, whereas in the colliding cue task no such relationship exists. The underlying similarity, then, appears to be the cognitive demands underlying attending to a particular feature in the context of interference from a second feature. As mentioned above, in early speech segmentation learners need to identify potential segmentation cues while ignoring salient cues that are not viable candidates. Likewise, the process of weighting cues as they are integrated may also contain a component of selective attention and/or inhibitory control.

One could argue, however, that given the integrative nature of speech segmentation, the methods employed in our study do not accurately depict the problems confronted by the language learner. Our study allows learners to select between two viable cues that indicate word boundaries at separate locations, whereas the process of speech segmentation involves partially overlapping cues that vary in reliability (see Christiansen et al., 2005). While we agree that the extent to which cues are consistent is likely an important determinant of how they are weighted and that integration may ultimately consist of finding the common ground across multiple cues (see Christiansen et al., 2005), we elected to adopt this paradigm for two reasons. Pragmatically, it made sense that our initial foray into this area should build on an existing paradigm. The colliding cue method used in developmental studies of speech segmentation (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003) was a logical choice. Furthermore, this methodology was also appropriate as a starting point since the paradigm represents a reductionist characterisation of the most challenging type of conflict that could occur in early language segmentation (in which the learner is confronted with two viable cues and must select one in order to successfully segment the stream). Given our interest in exploring the extent to which executive functioning might be related to resolving colliding cues in the speech stream, it made sense to increase the perceptual demands of the task in order to increase the likelihood of engaging these mechanisms (see below). Our future work will explore how this conception of cue competition can interface with recent models of word segmentation, such as the models proposed by Christiansen et al. and others.

That said, are learners ever confronted with two consistent cues to segmentation that occur at different locations in the speech stream? We argue it is possible. Penultimate stress provides an example of a consistent cue that is not at the word boundary. In an experiment by Toro-Soto and colleagues (2007), when learners initially confront an artificial speech stream that has both transitional probability cues and penultimate stress cues, learning drops to chance, even for Spanish speakers who use penultimate stress in their natural language. While that paper did not frame the results in terms of conflict, it is likely that the learners in that study defaulted to stress as a cue to word onset and thus had created a conflicting cue between stress and transitional probabilities. The authors do not report the individual data, but it would be of interest to know whether they obtained the same type of variance found in our study. While these cues are ultimately integrated by Spanish speakers, it stands to reason that they might not be initially integrated by infant learners (following the logic of Thiessen & Saffran, 2003) until they are able to stop perceiving stress as an indicator of word onset (a process that itself may have an inhibitory component).

In our study, the observed correlation failed to hold for all of the other experimental conditions in which the strength of one cue predominated over the other (as independently determined by performance on the segmentation test). Similarly, the control condition also yielded no correlation with the Simon test. There were two indications that the observed pattern in Condition 2 was not an epiphenomenon due to differences in variance across conditions. First, combining scores across all conditions (which increases variance) did not yield a significant correlation with performance on the Simon task. Second, the supplemental analysis comparing the slopes across conditions supported the notion that the observed pattern in the 50 ms condition was unique relative to the other experimental conditions. Rather, we argue from this pattern of results that a relationship exists between cognitive mechanisms that manage conflicting cues in general information processing tasks and the mechanisms necessary to resolve difficult colliding cue segmentation problems in language acquisition. This relationship is manifest when the perceptual demands of the task are highest, consistent with recent findings from visual processing suggesting that the degree of competition between two cues determines the extent to which top-down mechanisms are involved in resolving competition (see Torralbo & Beck, 2008). Increasing the number of items in a display increases the need for top-down biasing in order to identify a particular target. The same may be true for early speech segmentation in which many highly competitive cues interact, thereby increasing the demands for top-down processing.

This claim has important, and relatively unexplored, implications for our understanding of the early stages of language acquisition. While the last 10 years have provided data suggesting that domain general learning mechanisms play a role in the early stages of language acquisition (e.g., Saffran et al., 1996a), there have been few efforts to address whether the course of language acquisition is mediated or modulated by the development of general cognitive abilities. One such attempt to associate early mechanisms of language acquisition and more general cognition was an infant study examining general cognitive and perceptual abilities (Lalonde & Werker, 1995). Building on work demonstrating that at 6 months of age, infants' phonetic discrimination abilities are independent of their language experience, whereas by 12 months, infants are adult-like in that their phonetic discrimination abilities become limited by their native language input (Werker & Tees, 1984), Lalonde and Werker (1995) showed that changes in phonetic discrimination abilities coincided with changes in general cognitive abilities, as demonstrated by performance on object search and visual categorisation tasks. They concluded that age-related changes in language acquisition are linked to changes in other developing cognitive and perceptual skills. A more recent study by Conboy, Sommerville, and Kuhl (2008) further attested this assertion. In a study of 11-month-old infants, they found that nonnative speech discrimination (which typically declines toward the end of the first year of life) negatively correlated with measures of inhibitory control (indicating that the children who performed more similar to native adult speakers tended to score better on measures of inhibitory control). This finding suggests a relationship between inhibitory control and some of the early measures of language proficiency. Our data suggest that such a relationship may persist into adulthood and may include abilities such as selective attention or inhibitory control. These abilities are known to be available to infants in varying degrees. Selective attention has been demonstrated in very young infants, with abilities showing an increasing tendency over the first 6 months of life (Richards, 1998). Inhibitory control, which continues to develop well into adolescence, has roots that develop as early as 9 months of age (e.g., Amso & Johnson, 2005), roughly coinciding with the age at which infants appear to switch segmentation strategies (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003). The development of the dorsolateral prefrontal cortex during the first year of life is thought to support this increase in inhibitory control abilities during the first year of life (Diamond & Goldman-Rakic, 1989).

Our findings also attest to the flexibility of learners' speech segmentation abilities. Previous work on segmenting known words from a speech stream established that segmentation strategies may change as a function of background noise (e.g., Mattys et al., 2005) or cue strength (Mattys & Melhorn, 2007). Our results extend these findings to a speech segmentation task in which the learner has no *a priori* knowledge of the language, simulating the early stages of acquisition. We demonstrated that under these conditions, cue strength also plays a critical role in determining which segmentation strategy is deployed. In our control condition, learners heard an artificial language containing only transitional probability cues to word boundaries and performed significantly above chance levels. Across experimental conditions, incrementally increasing the length of competing pause cues increased the likelihood of parsing according to pause boundaries. In Condition 1, learners were more likely to segment the stream according to transitional probabilities, despite the insertion of 25 ms pauses. In Condition 2 learning was not significantly different from chance levels, suggesting that inserting pauses of 50 ms duration yielded a competing cue that was approximately equivalent in strength to the transitional probability cues. Moreover, closer investigation of the response distribution revealed that this condition exhibited the most variability, with some learners scoring high, others scoring low, and a sizeable group scoring at chance on the test. From this we observe that some learners were capable of parsing the stream using statistics (high scorers) or pauses (low scorers), while others failed to extract either of the consistent patterns (chance scorers). In Condition 3, the 75 ms pauses appeared to be robustly salient, and most learners parsed the stream accordingly. Thus, we conclude that the strength of a segmentation cue can directly impact the learner's choice of segmentation strategy. From this finding one can infer that variability in cue strength across languages could contribute to the development of segmentation strategies that differ from one language to the next (as evidenced, for example, in Cutler et al., 1986). Future developmental work employing colliding cue paradigms must be careful to account for this factor.

The larger implication of this body of work lends support to the notion asserted by Bates (1994) that language learning may be based on a relatively plastic combination of neural systems that also subserve other tasks. Minimally, our data suggest that individual differences in cognitive abilities may impact the course of language acquisition. In fact, this link provides a promising direction for enhancing our understanding of the fundamentally variable way in which people acquire language. Variability is one of the hallmarks of acquisition at all levels. For example, although all normally developing infants eventually acquire language, it is well known that the pacing and patterning of the developmental milestones varies from child to child (see Bates, Dale, & Thal, 1995). The problem is even more pronounced in late L2 learners, who exhibit tremendous variance with respect to ultimate proficiency (e.g., Johnson & Newport, 1989). While a multitude of factors likely underlie these observations (Bates et al., 1991), the results reported here, in conjunction with the findings of Lalonde and Werker (1995), point to individual differences in general cognitive abilities as an important source of explanatory power for the fundamentally variable nature of language acquisition.

Manuscript received June 2008
Revised manuscript received July 2009
First published online September 2009

REFERENCES

Amso, D., & Johnson, S. P. (2005). Selection and inhibition in infancy: Evidence from the spatial negative priming paradigm. *Cognition*, *95*(2), 27–36.

Bates, E. (1994). Modularity, domain specificity and the development of language. *Discussions in Neuroscience*, *1*(2), 136–149.

Bates, E., Bretherton, I., & Snyder, L. S. (1991). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge, UK: Cambridge University Press.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 96–151). Oxford, UK: Blackwell.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290–303.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Cairns, P., Shillcock, R., Chater, N., & Levy, J. (1997). Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology*, *33*(2), 111–153.

Christiansen, M. H., Conway, C. M., & Curtin, S. (2005). Multiple-cue integration in language acquisition: A connectionist model of speech segmentation and rule-like behavior. In J. W. Minett & W. S.-Y. Wang (Eds.), *Language acquisition, change and emergence: Essay in evolutionary linguistics* (pp. 205–249). Hong Kong: City University of Hong Kong Press.

Conboy, B. T., Sommerville, J. A., & Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, *44*, 1505–1512.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.

Diamond, A., & Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research*, *74*(1), 24–40.

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, *11*(6), 467–473.

Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, *67*(6), 3192–3218.

Grosjean, M., & Mordkoff, J. T. (2002). Post-response stimulation and the Simon effect: Further evidence of action-effect integration. *Visual Cognition*, *9*(4), 528–539.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 653–675). Cambridge, MA: MIT Press.

Hedge, A., & Marsh, N. W. A. (1975). The effect of irrelevant spatial correspondences on two-choice response-time. *Acta Psychologica*, *39*, 427–439.

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

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, *130*(4), 658–680.

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

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*, 279–312.

Lalonde, C. E., & Werker, J. F. (1995). Cognitive influences on cross-language speech perception in infancy. *Infant Behaviour and Development*, *18*(4), 459–475.

Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin and Review*, *2*(2), 174–207.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle.* Cambridge, MA: MIT Press.

Mattys, S. L. (2004). Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 397–408.

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language*, *47*(3), 343–359.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–494.

Mattys, S. L., & Melhorn, J. F. (2007). Sentential, lexical, and acoustic effects on the perception of word boundaries. *The Journal of the Acoustical Society of America*, *122*, 554–567.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General*, *134*(4), 477–500.

Mueller, J. L., Bahlman, J., & Friederici, A. D. (2008). The role of pause cues in language learning: The emergence of event-related potentials related to sequence processing. *Journal of Cognitive Neuroscience*, *20*(5), 1–14.

Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, *66*(4), 911–936.

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

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Newport, E. L., Weiss, D. J., Aslin, R. N., & Wonacott, E. (2004). *Statistical Learning in Speech: Syllables or Segments?* Paper presented at the 29th Annual Boston University Conference on Language Development, Boston, MA.

Peña, M., Bonatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven compuations in speech processing. *Science*, *298*, 604–607.

Perruchet, P., Tyler, M., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*, *133*, 573–583.

Richards, J. E. (1998). Development of selective attention in young infants: Enhancement and attenuation of startle reflex by attention. *Developmental Science*, *1*(1), 45–51.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Directions in Psychological Science*, *12*, 110–114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.

Simon, J. R., & Berbaum, K. (1990). Effect of conflicting cues on information processing: the 'Stroop effect' vs. the 'Simon effect'. *Acta Psychologica (Amsterdam)*, *73*(2), 159–170.

Simon, J. R., & Small, A. M. (1969). Processing auditory information: Interference from an irrelevant cue. *Journal of Applied Psychology*, *53*, 433–435.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- and 9- month-old infants. *Developmental Psychology*, *39*(4), 706–716.

Toro-Soto, J. M., Rodríguez-Fornells, A., & Sebastián-Gallés, N. (2007). Stress placement and word segmentation by Spanish speakers. *Psicológica*, *28*, 167–176.

Torralbo, A., & Beck, D. M. (2008). Perceptual load-induced selection as a result of local competitive interactions in visual cortex. *Psychological Science*, *19*(10), 1045–1050.

von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology*, *5*(4), 520–526.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development*, *5*, 30–49.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.