Testing the Limits of Statistical Learning for Word Segmentation

Elizabeth K. Johnson[1] and Michael D. Tyler[*]

[1] Department of Psychology, University of Toronto, Toronto, Canada

[*] School of Psychology and MARCS Auditory Laboratories, University of Western Sydney, Sydney, Australia

Address for Correspondence:

Elizabeth K. Johnson

3359 Mississauga Road N.

Room 4071, CCT

Mississauga, ON

Canada L5L 1C6

Abstract

Past research has demonstrated that infants can rapidly extract syllable distribution information from an artificial language and use this knowledge to infer likely word boundaries in speech. However, artificial languages are extremely simplified with respect to natural language. In this study, we ask whether infants' ability to track transitional probabilities between syllables in an artificial language can scale up to the challenge of natural language. We do so by testing both 5.5- and 8-month-olds' ability to segment an artificial language containing four words of uniform length (all CVCV) or four words of varying length (two CVCV, two CVCVCV). The transitional probability cues to word boundaries were held equal across the two languages. Both age groups segmented the language containing words of uniform length, demonstrating that even 5.5-month-olds are extremely sensitive to the conditional probabilities in their environment. However, neither age group succeeded in segmenting the language containing words of varying length, despite the fact that the transitional probability cues defining word boundaries were equally strong in the two languages. We conclude that infants' statistical learning abilities may not be as robust as earlier studies have suggested.

Keywords: artificial language, statistical learning, infant word segmentation

It has been suggested that infants initially learn to segment words from speech by tracking transitional probabilities between syllables (Thiessen & Saffran, 2003, 2007). Highly probable syllable transitions are perceived as likely within-word sequences whereas low-probability transitions are seen as likely between-word sequences (Aslin, Saffran, & Newport, 1998). Once a sufficient number of word forms have been segmented from speech using transitional probabilities between syllables, infants can begin learning other segmentation heuristics, such as the likelihood of stress signaling a word onset in English (see Saffran, Werker, & Werner, 2006, for review). According to this view, transitional probabilities between syllables are not the only cue to word boundaries; however, they are an extremely important cue in that their use enables infants to bootstrap all other speech cues to word boundaries. Such a statistical solution to the word segmentation problem is attractive because it provides a potentially language-general strategy for the acquisition of language-specific segmentation cues.

Proponents of this statistical bootstrapping theory typically cite two main lines of evidence. First, analyses of infant-directed corpora suggest that transitional probabilities between syllables provide good cues to word boundaries. For example, Swingley (2005) analyzed both Dutch and English infant-directed corpora, and concluded that syllable distribution cues could indeed enable infants to locate word boundaries. Second, 6.5- to 8-month-old infants can use transitional probabilities between syllables to rapidly segment words from an artificial language (Saffran, Newport, & Aslin, 1996; Thiessen & Saffran, 2003). Thus, infants possess extremely powerful statistical learning abilities, suggesting that the statistical bootstrapping theory is psychologically plausible. Moreover, infants appear to treat the syllable sequences that they extract from artificial languages as words (Graf Estes, Evans, Alibali, & Saffran, 2007; Saffran, 2001). For these reasons, the statistical bootstrapping theory has justifiably gained strong support among language acquisition researchers (Gómez, 2007; Saffran et al., 2006; Werker & Curtin, 2005).

Despite the evidence in favor of the statistical bootstrapping theory, there is still no direct evidence linking infants' ability to track transitional probabilities between syllables to the onset of infants' ability to segment words from every day natural speech. Moreover, there is good reason to question both of the abovementioned lines of evidence. First, although analyses of infant-directed corpora have demonstrated that transitional probabilities between syllables are a good indicator of word boundaries, these cues are undoubtedly stronger in the broadly transcribed speech from corpora than in actual fluent speech.[1] Second, although it is indisputable that infants possess powerful statistical learning mechanisms, the artificial languages that are used to demonstrate this are extremely simplified with respect to natural language, to the extent that they may introduce regularities that assist infants' detection of word boundaries (see Yang, 2004, for a related discussion).

In this study, we begin to ask whether infants' statistical learning abilities are robust enough to handle the complexity and variation that is characteristic of natural language. We address this question not by using natural language, but by using an artificial language that contains slightly more variation, and is thus slightly more natural, than the artificial languages used in earlier infant studies. There are many regularities in artificial languages, not present in natural language, that we could have focused on in the current study (e.g., syllable structure, word length, and token variability), but we chose to carry out the manipulation that we believed would result in the least extreme increase in task complexity: uniformity of word length. Infants have been shown to readily segment an artificial language consisting solely of trisyllabic words (Saffran et al., 1996; Johnson & Jusczyk, 2001) or disyllabic words (Thiessen & Saffran, 2003). How would infants perform with an artificial language containing a mixture of disyllabic and trisyllabic words? If infants' ability to track transitional probabilities between syllables provides the bootstapping required for learning additional language-specific word boundary cues, then one would expect this ability to be fairly robust. If removing the word-length regularity from artificial languages negatively impacts infants' ability to learn the statistical structure of the language, then this would indicate that infants can rapidly extract the statistical structure of an artificial language only if other aspects of that language are simplified to the point of being predictable. This in turn would raise the question of whether infants' solution to the word segmentation problem for artificial languages is qualitatively different from their solution for natural language.

As an additional side question, we also ask whether 5.5-month-olds are as skilled in tracking transitional probabilities between syllables as 8-month-olds. This question is important since infants begin segmenting words from speech by 6- to 7.5 months of age (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Jusczyk & Aslin, 1995), and presumably infants would have to be tracking transitional probabilities in the ambient language for some time before they would be able to use this information to infer word boundaries. The 5.5-month-olds we test here are one month younger than the youngest age group that has ever been tested in an artificial language segmentation task. If such young infants were to succeed at segmenting an artificial language using transitional probabilities alone, this would be an impressive demonstration of infants' early sensitivity to distributional regularities in the environment.

The artificial languages used in the current study contained four words composed of consonant-vowel (CV) syllables. In one condition, the artificial language contained four disyllabic words (henceforth, the uniform word length, UWL, condition). In the other condition, the artificial language contained two disyllabic words and two trisyllabic words (henceforth, the mixed word length, MWL, condition). In both conditions, the statistical structure of the language was the same as in Thiessen and Saffran (2003; transitional probabilities spanning word boundaries were equal to .33, whereas transitional probabilities between syllables within a word were equal to 1.0). All infants were exposed to an approximately 2.5-minute stream of speech, and then immediately tested for their listening preferences for words versus partwords. All test items were disyllabic. As in most artificial language studies testing this age group (e.g., Aslin et al, 1998; Curtin, Mintz, & Christiansen, 2005; Johnson & Jusczyk, 2001; Saffran et al., 1996; Saffran, Johnson, Aslin, & Newport, 1999), we predicted that infants would demonstrate their recognition of the words in the language by looking longer to novel (partword) test items. If both age groups succeed at segmenting both languages equally well, then this provides further support for the statistical bootstrapping hypothesis. If, however, either the 5.5-month-olds or both the 5.5- and 8-month-olds only succeed at segmenting the language containing words of uniform length, then this would suggest that infants' ability to track transitional probabilities between syllables is extremely easy to disrupt, and thus might not be robust enough to handle real language input.

Method

Participants

Dutch-learning infants from the Nijmegen region were recruited from two age groups: 5.5 and 8 months old. Forty-eight 8 month-olds (18 females) were tested, M_{age} = 265 days, Range = 258 to 275 days, and the data from 14 additional infants was excluded due to fussiness (7), falling asleep (1), computer error (3), parental/sibling interference (2), or looking times greater than 3 *SD* from the mean (1). Forty-eight 5.5-month-olds (23 females) were also tested, M_{age} = 179 days, Range = 163 to 197 days, with the data from 9 additional infants excluded due to fussiness (4), refusal to look at lights (4), or parental/sibling interference (1). Parental consent was obtained for all participants.

Stimuli

The artificial languages consisted of the concatenation of four disyllabic words (UWL condition), or two disyllabic and two trisyllabic words (MWL condition), and were synthesized using MBROLA's (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996) female nl3 diphone set. The CV syllables came from a pool of six consonants (two stops, /b/, /k/; two fricatives, /f/, /s/; a nasal, /m/, and; a liquid, /l/), and four vowels (/iː/, /ɛː/, /ɔː/, and /uː/). All consonants were 100 ms long and all vowels were 173 ms long, resulting in a total syllable length of 273 ms, which is the same as that used by Johnson & Jusczyk (2001). The fundamental frequency was set to a monotone 220 Hz.

The familiarization language was 2 min 27 s long in the UWL condition and 3 min 17 s long in the MWL condition. A 5-s fade-in and fade-out was applied to the beginning and end of the familiarization streams to avoid word boundary cues. To overcome potential native-language word boundary biases (see Reber & Perruchet, 2003), each infant was presented with a different artificial language; that is, a different pseudorandom combination of consonants and vowels was used to create the words of the language. The six consonants and four vowels were randomized and then allocated to words in a fixed pattern, such that each word and test partword started with a different consonant, and each vowel occurred only once in a given syllable across the four words of the language. The words were concatenated randomly to create the familiarization stream, with the constraint that the same word could not immediately follow itself. Note that the end of one word was fully coarticulated with the beginning of the following word.

Two test partwords were constructed by combining the last syllable of one word with the first syllable of another word. All test items in both conditions were disyllabic. To ensure that the two test words and the partwords had identical word frequencies (see Aslin et al., 1998), the words used to construct the partwords were presented in the familiarization stream twice as often (90 times) as the other two words that served as test items (45 times).

Procedure and apparatus

Infants were tested using the same version of the Headturn Preference Procedure used by Saffran et al. (1996). Infants sat on their caregiver's lap in the center of a threesided booth. The experimenter viewed the infant on a monitor, and recorded the infants' head movements via a button box. A blue light and a speaker were mounted at eye level on the center of each side panel, and a red light was located on the center of the front panel. During the familiarization phase, the speech stream played from both speakers and the lights in the booth were lit and extinguished contingent upon the infant's headturns.

The test phase followed immediately after the familiarization phase. Each of the 12 test trials (three trials for each of the two words and two partwords) began with the blinking center light. Once the infant oriented forward, the light stopped blinking and one of the two side lights began blinking. Once the infant oriented toward the blinking light, a test item was repeated from the speaker behind the flashing light with a 500 ms ISI until the infant looked away for more than 2 consecutive seconds or until 15 repetitions of the test item had occurred. Test trials were blocked and presented in random order within those blocks. To avoid bias, the caregiver and experimenter listened to masking music over Aviator-style headphones. The dependent measure in this study was orientation times to test stimuli.

Results and Discussion

Mean orientation times for both age groups, and both language type conditions, are presented in Figure 1. In the UWL condition, 16 of the 24 5.5 month-olds and 16 of the 24 8 month-olds oriented longer to partwords than to words. The mean orientation times for 5.5 month-olds were 7.44 s, *SE* = 0.42, for words and 8.38 s, *SE* = 0.49 for partwords, and for the 8 month-olds they were 6.56 s, *SE* = 0.37, and 7.64 s, *SE* = 0.41, respectively. In the MWL condition, 12 of the 5.5 month-olds and 14 of the 8 month-olds oriented longer to partwords than to words. The 5.5-month-olds' mean orientation times were 7.14 s, *SE* = 0.52, for words and 7.03 s, *SE* = 0.56, for partwords, and for the 8 month-olds they were 8.13 s, *SE* = 0.53, and 8.31s, *SE* = 0.50, respectively.

INSERT FIGURE 1 ABOUT HERE

The mean orientation times were subjected to a 2 (Age) x 2 (Language Type) x (2) (Test Item Type) ANOVA, with repeated measures on the third factor, with an alpha rate of .05. There was no main effect of age, $F(1,92) = 0.14$, or language type, $F(1, 92) = 0.12$, but there was a significant main effect of test item type, $F(1,92) = 4.93$, $\eta_p^2 = .05$, such that orientation times were longer overall for partwords than words. Importantly, there was a significant two-way interaction between language type and test item, $F(1,92) = 8.62$, $\eta_p^2 = .09$, but no three-way interaction between those two factors and age, $F(1,92) = 0.07$. A simple effects test (Winer, Brown, & Michels, 1991) showed infants, collapsed across age, oriented longer to partwords than words in the UWL condition ($F(1,92) = 13.30$, $\eta_p^2 = .13$), but not in the MWL condition ($F(1,92) = 0.26$).

From these results, it is clear that both 5.5- and 8-month-old infants learned to segment the words from the artificial language when the words were all of the same length (the UWL condition). Note that this represents the first evidence that infants under 6 months of age can track transitional probabilities between syllables, and it is the first demonstration that infants learning a language other than English use transitional probabilities to segment an artificial language. On their own, these findings add credence to the claim that infants rely on transitional probabilities between syllables to extract their first words from speech. However, when words were of varying lengths (the MWL condition), neither age group succeeded in segmenting the language. The removal of the regularity that all words had the same length appears, therefore, to have seriously hindered infants' ability to segment the language. These results suggest that earlier studies using highly simplified input may have overestimated infants' ability to track transitional probabilities in continuous speech. When the additional regularity of uniform word length was removed, infants failed to show any evidence of segmenting the language. This suggests that tracking transitional probabilities between syllables might not be the most likely explanation for how infants first begin to solve the word segmentation problem.

One possible confound is that the languages used in the UWL condition had two more syllables than the languages used in the MWL condition. Perhaps this difference led to our failure to find an effect the MWL condition because it is easier to segment a language containing 8 syllables than one containing 10 syllables, or the slightly longer familiarization in the MWL condition may have led infants to be fussier and fail to show any looking time differences. Both of these alternatives are highly unlikely. First, there have been many studies demonstrating that infants readily extract the statistical structure of artificial languages containing 12 syllables (four trisyllabic words; Aslin et al., 1998; Johnson & Jusczyk, 2001; Saffran et al., 1996; Saffran et al., 1999). Second, in a pilot study carried out with English-learning 8-month-olds (Johnson & Jusczyk, 2003), the same pattern of results was obtained when infants were tested on a mixed-length language containing the same number of syllables as the uniform word length language.

Lastly, we doubt that the familiarization duration was problematic in the MWL Condition because past artificial language segmentation studies have succeeded in showing successful segmentation with a 3-min familiarization period (e.g. Aslin et al, 1998). Moreover, the experimenters, who were blind to the purposes of the experiment, reported that the infants run in MWL Condition were not exceptionally fussy compared to those run in UWL Condition and, furthermore, the overall orientation times were no shorter in the MWL than the UWL condition. In fact, there was a trend towards longer orientation times in the MWL condition so the longer exposure phase does not appear to have caused the infants' to lose interest in the stimuli. In short, we find the uniform versus mixed word length difference between the two conditions as the only reasonable explanation for our results.

General Discussion

Infants' ability to segment words from an artificial language containing no cues to word boundaries other than the transitional probabilities between syllables is often cited as support for the theory that infants use transitional probabilities to solve the word segmentation problem. For example, it has been suggested that infants learn which acoustic characteristics define word boundaries only after they begin segmenting words from speech based on transitional probabilities between syllables (Thiessen & Saffran, 2003, 2007). We do not deny the importance of distributional learning in language acquisition; however, we question the notion that transitional probabilities between syllables necessarily enable infants to initially begin segmenting words from speech. In the introduction, we cited two major weaknesses in the statistical bootstrapping theory for word segmentation. First, evidence that transitional probabilities define word boundaries in an infant-directed corpus does not demonstrate that 5- to 6-month-olds have the phonological knowledge or processing capabilities necessary to extract these regularities. Second, it is not clear that the statistical learning abilities infants demonstrate when faced with an artificial language can scale up to handle real language input. In this paper, we begin to address both issues. Our results indicate that both 5.5- and 8-month-olds have the processing capacity to extract statistical regularities from a highly simplified language, however, this ability appears to break down when the input language is even slightly more complex than that used in earlier studies. Thus, our results suggest that infants' ability to track transitional probabilities between syllables is much more fragile than earlier studies have suggested. This in turn suggests that an alternative, or at least revised, explanation may be needed to describe how infants initially solve the word segmentation problem.

To reiterate, our findings do not refute the ability of infants to track transitional probabilities. Indeed, we know of no other explanation for how infants succeeded in our UWL condition. Our point is simply that natural languages vary a great deal from artificial languages, and tracking transitional probabilities might be a segmentation solution better suited for artificial than natural languages. We do not see our results as indicating that infants are incapable of segmenting words from the MWL. If we had lengthened the familiarization phase, infants may have succeeded. Likewise, if we had replaced the word length regularity with another type of regularity (e.g. all words start with a labial consonant), we may have observed successful segmentation of the mixed word length language. Indeed, Thiessen, Hill, & Saffran (2005) report that 6.5- to 7.5-month-olds learn an artificial language containing words of varying length when the language is produced with infant- but not adult-directed prosody. This finding is perfectly consistent with the current findings. As an artificial language becomes more complicated (e.g. containing words of varying as opposed to uniform length), other natural speech cues such as exaggerated prosody are needed to make word segmentation possible. Here we removed one regularity, unrelated to transitional probabilities, that is typically present in the artificial languages used with infants and they no longer showed evidence of segmentation. This raises questions as to how strongly infants rely on this information when segmenting natural language.

The goal of the current study was not to downplay the impressive nature of infants' statistical learning abilities. Clearly, infants are highly skilled at detecting probabilistic patterns in their input. Rather, we view this work as a small step towards beginning to critically examine the psychological plausibility of statistical bootstrapping theories of word segmentation. One might criticize that we have complicated the segmentation task, and yet we still expect infants to extract words in a very brief period of time. But, in fact, we are still presenting infants with a language that is highly simplified with respect to natural language; the transitional probabilities defining within- versus between-word sequences are as exaggerated as those in earlier artificial language studies, and listening to the same four words repeating one after another at a constant speech rate continuously for 2.5 minutes would never happen in a real language learning setting (see Johnson & Seidl, 2009, for discussion). The fact that infants were only successful at segmenting the uniform word length language suggests that the rhythmic word length regularity of this language may have made it easier to learn. This factor was not considered in earlier artificial language studies of this type. Note that the major contribution of Saffran et al. was to show that infants could segment an artificial language containing no cues to word boundaries other than the transitional probabilities between syllables. What our results suggest is that this may not be the case. The infants tested by Saffran et al. my have succeeded due to the presence of not only transitional probability cues, but also word length regularities.[2] These regularities do not exist in natural language. To justify claims that infants initially solve the word segmentation problem in natural language by tracking transitional probabilities, further studies examining these types of issues are required. This sort of research is integral to understanding the role transitional probabilities between syllables play in natural language acquisition.

If future studies continue to show that statistical learning does not scale up to natural language, and tracking transitional probabilities between syllables cannot explain how infants begin extracting words from speech, then what might explain infants' eventual segmentation of their native language? Some have argued that words in isolation play a major role in helping infants bootstrap their way into word segmentation (Brent, 1999; Brent & Siskind, 2001; Peters, 1983; Pinker, 1984). Others have argued that highly frequent words, such as names and function words, are of paramount importance (Bortfeld et al., 2005; Shi, Cutler, Werker, & Cruickshank, 2006). Still others have argued that prosody is essential to explaining early word segmentation (Johnson, 2008; Johnson & Jusczyk, 2001; Jusczyk, 1997; Johnson & Seidl, 2008, 2009; Shukla, Nespor,

& Mehler, 2006). Clearly, there are many possible solutions to the word segmentation problem. The theory we favor most for explaining early word segmentation is prosodic in nature. Here we use the term ‘prosodic’ not to refer to the lexical stress cues often discussed in the infant artificial language learning literature (e.g. Johnson & Jusczyk, 2001; Johnson & Seidl, in press; Thiessen & Saffran, 2003; 2007), but to utterance-level prosody. Clause and phrase boundaries not only necessarily coincide with word boundaries, they are also highly salient to infants (Gout, Christophe, & Morgan, 2004; Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000; Soderstrom, Kemler Nelson, & Jusczyk, 2005). Infants could begin to learn segmentation strategies by attending to the ends of these units. It is possible that language-specific segmentation strategies, such as the Metrical Segmentation Strategy (Cutler & Butterfield, 1992), could be learned by attending to the edges of prosodically marked clauses and phrases.[3] Support for this idea is provided by infants’ tendency to segment words occurring along utterance edges more readily than words that are not flanked by strong prosodic boundaries (Seidl & Johnson, 2006; 2008). Note that this prosodic explanation for early word segmentation can also account for reports that isolated words as well as names and function words play an important role in early word segmentation. Isolated words are flanked by strong prosodic boundaries, and both names and function words are not only highly frequent, but they also tend to be aligned with a prosodic boundary. In short, tracking transitional probabilities between syllables is not the only viable solution to the word segmentation problem. Indeed, no one has ever claimed that infants use transitional probabilities alone to segment words from language. It has simply been suggested that transitional probabilities might be the first cue infants use to segment speech (Thiessen & Saffran?).

We would like to suggest that transitional probabilities may possibly contribute towards infants' segmentation abilities, but we find it unlikely that sensitivity to this statistical cue is the driving force that enables infants to first begin segmenting words from speech. Certainly the findings reported in this paper suggest that infants' ability to track transitional probabilities might not be as robust as past studies have suggested. Only future research will enable us to determine whether the statistical bootstrapping theory or some other proposal provides the best explanation for how infants begin to segment words from speech.

References

Aslin, R.N., Saffran, J.R., & Newport, E.L. (1998). Computation of conditional probability statistics by human infants. *Psychological Science*, *9*, 321-324.

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

Brent, M.R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Science*, *3*, 294-300.

Brent, M., & Siskind, J. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*, B33-44.

Curtin, S., Mintz, T.H., & Christiansen, M.H. (2005). Stress Changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*, 233-262.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*, 218-236.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vrecken, O. (1996). The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393-1396).

Gómez R.L. (2007). Statistical learning in infant language development. In M. Gareth Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 601-615). New York: Oxford University Press.

Gout, A., Christophe, A, & Morgan, J.L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, *51*, 548-567.

Graf Estes, K., Evans, J.L., Alibali, M.W., & Saffran, J.R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*, 254-260.

Johnson, E.K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *Journal of the Acoustical Society of America*, *123*, EL144-EL148.

Johnson, E. K., & Jusczyk, P.W. (2003). Exploring statistical learning by 8-month-olds: The role of complexity and variation. In D. Houston, A Seidl, G.Hollich, E. Johnson, & A. Jusczyk (Eds.) *Jusczyk Lab Final Report* (pp. 141-148).

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, E.K., & Seidl, A. (2008). Clause segmentation by 6-month-olds: a crosslinguistic perspective. *Infancy*, *13*, 440-455.

Johnson, E.K., & Seidl, A. (2009). At 11 months, prosody still outranks statistics. *Developmental Science*, *12*, 131-141.

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

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

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

Nazzi, T., Kemler Nelson, D.G., Jusczyk, P.W., & Jusczyk, A.M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. *Infancy*, *1*, 123-147.

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

Pinker, S. (1984). *Language Learnability and Language Development*. Cognitive Science Series. Cambridge, MA: Harvard University Press.

Reber, R., & Perruchet, P. (2003). The use of control groups in artificial grammar learning. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *56A*, 97-115.

Saffran, J.R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149-169.

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

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

Saffran, J.R., Werker, J.F., & Werner, L.A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of Child Development* (6th ed.) (pp. 58-108). New York: Wiley.

Seidl, A. & Johnson, E.K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science*, 9, 566-574.

Seidl, A. and Johnson, E.K. (2008). Perceptual factors influence infants' extraction of onsetless words from continuous speech. *Journal of Child Language*, *35*, 1-24.

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*, EL61-EL67.

Shockey, L. (2003). *Sound patterns of spoken English*. Oxford: Blackwell Publishing.

Shockey, L., & Bond, Z.S. (1980). Phonological processes in speech addressed to children. *Phonetica*, *37*, 267-274.

Shukla, M., Nespor, M., & Mehler, J. (2006). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, *54*, 1-32.

Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, *50*, 86-132.

Soderstrom, M., Kemler Nelson, D.G., & Jusczyk, P.W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development*, *28*, 87-94.

Thiessen, E.D., Hill, E., & Saffran, J.R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*, 53-71.

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

Thiessen, E.D., & Saffran, J.R. (2007). Learning to learn: Infants' acquisition of stress-based segmentation strategies for word segmentation. *Language, Learning and Development*, *3*, 73-100.

Werker, J.F., & Curtin, S. (2005). PRIMIR: A developmental model of speech processing. *Language Learning and Development*, *1*, 197-234.

Werker, J.F., & Tees, R.C. (1999). Experiential influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, *50*, 509-535.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Yang, C.D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, *8*, 451-456.

Acknowledgments

This research was supported by the MPI and an NWO Spinoza Prize awarded to Anne Cutler. Michael Tyler was also supported by a visiting fellowship from the Max Planck Institute for Psycholinguistics and NIHCD grant DC00403 (PI: C. Best). We thank the families who took part in this research, as well as Angela Khadar and Margret van Beuningen for assistance in recruiting and running participants. We would also like to credit Peter Jusczyk for inspiring this study. These results were previously presented at ISIS 2006 (Kyoto, Japan), AMLAP 2006 (Nijmegen, The Netherlands), and ESPP 2008 (Utrecht, The Netherlands).

Footnotes

[1] Connected speech processes obscure regularities found in transcribed speech (Shockey, 2003; Shockey & Bond, 1980), and we know little about infants' ability to cope with this. Furthermore, 6- to 8-month-olds have not yet acquired their language's consonant inventory (Werker & Tees, 1999). Thus, any computations infants performed over their input would vary greatly from an analysis of transcribed speech.

[2] It may be the case that, in addition to learning the transitional probabilities between syllables, infants in the uniform word length language may have learned that all words were three syllables in duration. If so, this may have acted as an additional cue used by infants to detect likely word-boundary locations (see Thiessen & Saffran, 2007, for a similar discussion regarding lexical stress and transitional probabilities). If this is the case, then infants' success at segmenting the language was not only due to their sensitivity to transitional probability cues. It was due to their sensitivity to transitional probabilities between syllables plus word length regularities.

[3] Swingley (2005) demonstrates that the majority of disyllabic utterances heard by infants are not stress-initial. However, the vast majority of short utterances in general begin with a stressed syllable. Thus, infants may possibly learn that stress signals the onsets of words by attending to utterance edges. We thank Charles Yang for this observation.

Figure Caption

Figure 1. Mean orientation times for 5.5- and 8-month-olds in the uniform word length and mixed word length conditions. Error bars represent standard error of the mean.