Effects of simultaneously presented visual information on adults' and infants' auditory statistical learning

Erik D. Thiessen
Department of Psychology
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh PA 15213

Abstract

Infant and adult learners are able to identify word boundaries in fluent speech using statistical information. Similarly, learners are able to use statistical information to identify word-object associations. Successful language learning requires both feats. In this series of experiments, we presented adults and infants with audio-visual input from which it was possible to identify both word boundaries and word-object relations. Adult learners were able to identify both kinds of statistical relations from the same input. Moreover, their learning was actually facilitated by the presence of two simultaneously present relations. Eight-month-old infants, however, do not appear to benefit from the presence of regular relations between words and object. Adults, like 8-month-olds, did not benefit from regular audio-visual correspondences when they were tested with tones, rather than linguistic input. These differences in learning outcomes across age and input suggest that both developmental and stimulus-based constraints affect statistical learning.

Keywords: statistical learning, cross-modal stimuli, development of cross-modal integration

Introduction

Learners are able to identify many different kinds of statistical regularities from linguistic input, including phonological and syntactic patterns (e.g., Chambers, Onishi, & Fisher, 2003; Mintz, 2002; Thiessen & Saffran, 2003). Despite the power of statistical learning, though, there is little doubt that human learners are constrained. Learners do not identify all kinds of statistical patterns equally well (e.g., Newport & Aslin, 2004; Peperkamp, le Calvez, Nadal, & Dupoux, 2006; Redford, 2008; Saffran & Thiessen, 2003). However, most of the research into constraints on human learning has focused on how learners do when presented with a single learning task. This is insufficient for a complete understanding of statistical learning for two reasons. First, language frequently presents learners with multiple problems simultaneously. For example, when exposed to a novel word form in fluent speech, learners have the opportunity to both learn the word form, and to learn the referent of the word. Second, constraints on learning may be especially important when the input is complex enough to support multiple learning problems (e.g., Fiser & Aslin, 2002; Pinker, 1984).

Consider the interaction between the statistical information useful for segmenting words from fluent speech (e.g., Saffran, Aslin, & Newport, 1996), and identifying referents for words (e.g., Smith & Yu, 2008). Taken in isolation, both word segmentation (e.g., Thiessen, Hill, & Saffran, 2005; Toro, Sinnett, & Soto-Faraco, 2005) and referential learning are constrained (Golinkoff, Shuff-Bailey, Olguin, & Ruan, 1995; Landau, Smith, & Jones, 1988; Markman, 1990; Markman & Wachtel, 1988). It is also clear that these learning processes interact. Learners who are previously familiar with a word form map it more easily to a novel referent (e.g., Graf Estes, Evans, Alibali, and Saffran, 2007; Storkel et al., 2001). Conversely, children map familiar objects to novel labels more easily than unfamiliar objects (e.g., Hall, 1991). Because these learning tasks interact, different constraints may operate when learners are presented with both problems simultaneously. If the interaction between the problems is unconstrained, the additional complexity when they are presented together may hinder learning (Fiser & Aslin, 2002; Pinker, 1984). Alternatively, learning may be constrained in such a way that the learning occurs sequentially, with one problem privileged and learned first. It is even possible that learning, if appropriately constrained, could be facilitated by the simultaneous presentation of multiple regularities. This could be the case if learning of one regularity reinforces the other.

To explore these possibilities, it is critical to present learners with the opportunity to identify simultaneous regularities. This set of experiments did so by building on prior research demonstrating that learners benefit from the embedding of audio input in a visual context (e.g., Hollich, Newman & Jusczyk, 2005). Appropriate visual information helps learners determine whether speakers are producing one language or multiple languages (Soto-Farco et al., 2007). Similarly, the presence of a video improves adults' ability to identify word boundaries in fluent speech (Sell & Kaschak, under review). In all of these tasks, however, the auditory learning task is the only task, and vision facilitates that task. The current experiments differ by presenting learners with two problems simultaneously: word segmentation, and discovery of word-object relations. This better simulates the richness of language, where any single utterance may provide information about many different aspects of language (e.g., Saffran & Wilson, 2001).

Experiment 1

All learners in this experiment were presented with words embedded in fluent speech. As in previous statistical learning experiments (e.g., Saffran et al., 1996), these words could be segmented via use of transitional probabilities that were high within words, and low at word boundaries. A subset of the participants in this experiment (in the *no-video* condition) were presented solely with fluent speech. The only learning task this group faced was identifying word boundaries.

A second group of participants (in the *regular-video* condition) saw objects synchronized to the onset and offset of the words in the fluent speech. Each word in the fluent speech was consistently paired with a unique object. As such, this group of participants was presented with two potential statistical regularities to learn: word boundaries, and the relations between particular words and objects.

A third group of participants (in the *irregular-video* condition) also saw shapes synchronized to words in fluent speech, but these participants saw objects that were not consistently associated with the words. This condition serves as a control to make sure that performance in the regular-video condition is not affected by some aspect of the visual stimuli other than the regular relation between words and shapes.

Method

Participants

Participants were 60 undergraduates at Carnegie Mellon University. Twenty participants apiece were randomly assigned to one of three stimulus conditions: no-video, irregular-video, or regular-video.

Stimuli

Audio Stimuli

All participants were exposed to a stream of synthesized speech used in Saffran et al.'s (1996) experiments. This artificial language contained four words: *padoti*, *bidaku*, *tupiro*, and *golabu*. The transitional probabilities between syllables within a word were 1.0, and the transitional probabilities between syllables across word boundaries were .33. Two words (*bidaku* and *tupiro*) and two part-words (*tigola* and *bupado*) were used as test items. Unlike words, part-word test items contained a transition between syllables with low transitional probability.

Visual Stimuli

In the *no-video* condition, participants saw a static checkerboard image for the duration of their exposure to the synthesized speech.

Participants in the *regular-video* and *irregular-video* condition saw looming shapes synchronized with the word boundaries. Shapes appeared at the same instant the word began to play, and remained onscreen for the duration of the word. At the beginning of a word, each shape occupied roughly $1/16^{th}$ of the screen. Over the course of the presentation of the word, the shape increased in size until it filled the screen.

In the regular-video condition, each word was paired with a particular object (*padoti*: white cross; *bidaku*: green diamond; *tupiro*: purple heart; *golabu*: yellow hexagon). In the irregular-video condition, words and shapes co-occurred with no consistent pattern. *Procedure*

In all three conditions, participants sat in front of a portable DVD player with a 10'' screen wearing airline-pilot style headphones. Participants were simply informed that after watching the video, they would answer a series of questions about what they saw and heard.

Segmentation Test

There were 16 two alternative forced choice questions in the segmentation test. For each question, participants heard a word and a part word (in counterbalanced order), separated by one second of silence. They were asked to circle the item that sounded more like the speech they heard (for discussion of this procedure, see Saffran et al., 1997).

Word-Shape Correspondence test

After completion of the 16 segmentation test items, participants in only the regular-video condition were informed that they would now answer an additional series of 16 questions. These questions assessed whether participants learned that particular words corresponded to shapes. For each question, participants heard one of the four words from the synthesized speech. They then saw a sequence of four shapes on the screen, looming with the same animation as during the initial exposure. They were asked to circle which of the four shapes went with the word.

Results

A one-way ANOVA was performed on participants' scores on the word-segmentation test as a function of condition. There was a significant effect of condition, $F(2,57) = 5.4, p < .01$. Participants performed best in the regular-video condition ($M = 12.0, SE = 0.5$), and less well in the irregular-video ($M = 9.9, SE = 0.5$) and no-video condition ($M = 8.9, SE = 0.5$). Scores in all three condition differed from chance (all condition: binomial $p < .05$). To follow up the effect of condition indicated by the ANOVA, planned t-tests were performed. Here and elsewhere, all t-tests reported are two-tailed. There was no significant difference between participants' performance in the no-video and in the irregular-video condition: $t(38) = 1.1, p = .30$. However, participants in the regular condition scored significantly better than participants in either of the other two conditions (regular- vs. no-video: $t(38) = 3.4, p < .01$; regular- vs. irregular-video: $t(38) = 2.2, p < .05$).

Participants in the regular-video condition also learned word-object relations. On average, participants scored 8.8 (out of 16; chance = 4) correct on the correspondence test ($SE = 0.8$), which was significantly above chance, binomial $p < .01$. Further, as illustrated by Figure 2, the correlation between the two tests was positive, $r = .64$, and significant, $p < .01$. Higher scores on one test were associated with higher scores on the other. Results from the segmentation and correspondence test converge to indicate that the presence of regular word-object relations facilitated learning

Experiment 2

Prior experiments have demonstrated that infants are able to segment words from fluent speech via transitional probabilities (e.g., Saffran et al., 1996), and identify relations between words and shapes (e.g., Thiessen, 2007), but no experiments have assessed both simultaneously. Because infants are the primary learners of language, their performance is both theoretically and pragmatically important. For example, given the capacity limitations of infants, it is plausible to hypothesize that they would fail to integrate audio and visual information as effectively as adults. If so, they may not benefit from the audio-visual corresponded in the regular-video condition.

Method

Participants

Participants in this experiment were 60 infants between the ages of 7.5 and 9 months ($M = 8.26$). Infants were randomly assigned to one of three groups: no-video, regular-video, and irregular-video. In order to obtain data from 60 infants, it was necessary to test 66. The additional six infants were excluded for the following reasons: fussing or crying (3), parental interference (2), and experimenter error (1). According to parental report, all infants were full term, and free of ear infections at the time of testing.

Procedure

This experiment used a slightly modified version of the HPP, presenting the visual stimuli on a central monitor rather than from the side of the room. Preferential looking experiments with a central monitor are commonly and successfully used with infants (e.g., Fernald, 1985). Infant participants were seated on their parents' lap in a sound-isolated room, approximately one foot away from a 30'' monitor. There were two speakers adjacent to the monitor and a camera mounted above it. The parents wore noise-canceling headphones to eliminate bias. An experimenter outside the room watched the infant over a closed-circuit monitor to initiate test trials and code the direction of the infants' gaze.

There were two phases to this experiment: the segmentation phase, and the test phase. During the segmentation phase, infants heard the synthesized speech from speakers adjacent to the monitor, while the monitor displayed the visual stimuli appropriate to the infants' condition.

The test phase used the same two words and two part-words as the adult test. Each item was repeated 3 times, for a total of 12 trials. Before each trial, an attention-getter (a brightly colored Winnie the Pooh video, coupled with an excited exclamation) attracted infants' gaze to the monitor. Once the infant oriented to the monitor, the experimenter initiated the test trial. Each trial consisted of a repetition of a single word (or part-word), with a pause of 1 second between repetitions. For as long as infants' gazed at the monitor, the test item continued to repeat. When infants looked away from the monitor for two continuous seconds, the test trial ended.

Stimuli

The stimuli during the segmentation phase were presented for 50 seconds and were identical to the audio and video presentations used in Experiment 1. This exposure is half of the length in Experiment 1; pilot testing indicated that 100 seconds yielded an unacceptably high fuss-out rate. The test items were also identical to Experiment 1. During test phase, words and part-words were paired with an orange bar rotating like a propeller (it completed one revolution every three seconds). Pilot testing indicated that infants were far more likely to maintain their interest in the experiment if the monitor displayed a moving object rather than a static image. Both the color and the shape of the bar were novel with respect to the segmentation phase of the experiment, and the motion was unlike the looming animation infants saw during the segmentation phase.

Results

Infants in the no-video condition looked at word trials for 12.4 sec ($SE = 1.0$), and at part-word trials for 11.8 sec ($SE = 0.9$). This difference in looking trials between words and part-words was not significant, $t(19) = 1.1, p = .28$. Infants in the regular-video condition looked at word test trials for 9.7 sec ($SE = 0.7$), and to part-word test trials for 11.7 sec ($SE = 0.8$). This difference was significant, $t(19) = 3.7, p < .05$. Infants in the irregular-video condition showed the same pattern, looking at words ($M = 9.9$, $SE = 1.2$) less than part-words ($M = 11.6$, $SE = 1.1$). The difference in looking time to words and part-words was also significant for infants in this group: $t(19) = 2.5, p < .05$. Unlike infants in the no-video condition, infants in both the regular- and irregular-video condition listened longer to part-words than to words. This indicates that they had learned enough about the identity of words to distinguish them from part-words.

Infants appeared to perform better in the regular-video condition than in the no-video condition, as infants in the no-video condition failed to respond differentially to word and part-word trials. However, infants' performance in the regular- and irregular-video conditions was not significantly different, as indicated by a 2 (condition) x 2 (test item) ANOVA. As expected, since infants in both groups showed a preference for part-words, there was a main effect of test item: $F(1, 38) = 12.6, p < .01$. There was no main effect of condition: $F(1, 38) < 1$. There was also no interaction between test item and condition: $F(1, 38) < 1$. That is, infants' preference for part-words in the irregular-video condition was statistically equivalent to infants' preference in the regular-video condition. These analyses indicate that while infants may benefit from the presence of looming shapes synchronized with word boundaries (present in both video conditions), they do not gain an added advantage from the regular relations between words and shapes present in the regular-video condition. These results present two compelling questions, discussed separately below.

Why do infants fail to distinguish between words and part-words in the no-video condition, when they can do so in the regular- and irregular-video condition?

One possible explanation is that infants in the regular- and irregular-video conditions received some benefit not present in the no-video condition. The looming shapes may have facilitated learning by maintaining infants' attention (e.g., Frick & Richards, 2001; Thiessen, et al., 2005). Another possible benefit that infants may have received in both the regular- and irregular-video condition is the synchronization between the appearance of the shapes and word boundaries. For young infants, synchronization is one of the most important factors that enable identifying links between audio and visual events (e.g., Bahrick, Flom, & Lickliter, 2002; Gogate & Bahrick, 1998; Lewcowicz, 1986; 2003). Infants may have relied upon synchronization as a cue to word boundaries, a cue that was equally available in both the regular- and irregular-video conditions.

Why do infants, unlike adults, fail to benefit from the regular relations between words and shapes available in the regular-video condition?

The fact that infants' performance in the irregular-video condition is equivalent to their performance in the regular-video condition suggests that infants failed to detect the relations between words and shapes present in the regular-video condition. This suggestion is consistent with a variety of converging evidence indicating that infants at this age are relatively insensitive to relations between words and objects in the visual world. Eight-month-old infants have a small vocabulary (e.g., Fenson et al., 2002). In controlled word-learning experiments, infants typically fail to acquire names for novel objects until around a year of age (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). If infants cannot detect the relation between words and objects in the regular-video condition, they cannot benefit from any facilitation that identifying the relation provides to adult learners.

Experiment 3

There are at least two (not mutually exclusive) factors that can explain why infants in Experiment 2 failed to benefit from the regular audio-visual pairing, unlike adults in Experiment 1. One is that adults' ability to take advantage of the regular-video condition is due to the fact that they are faster, more efficient information processors than infants (e.g., Pelphrey & Reznick, 2003). To detect a relation between words and shapes, learners must process the identity of the shape (and the word) in a brief time. There are several experimental results suggesting that young infants are less successful in processing multiple sources of information than older infants and adults (e.g., Stager & Werker, 1997). A second potential factor is that the difference between 8-month-olds and older learners is due to differences in their prior linguistic experience. Adults are well aware that one of the primary functions of words is to refer to features of the visual world such as shape. Eight-month-olds may not yet expect to discover relations between words and shapes (cf. Werker et al., 1998). Infants may fail to detect the regular relations in the input because they do not expect them.

Both factors converge to suggest that older infants should be more successful in identifying and benefiting from regular word-object associations. Thus, in Experiment 3, we presented 20-month-old infants with the same stimuli used in Experiment 2. These children have a year of additional word-learning experience, and more advanced cognitive processing abilities. Should infants of this age fail to benefit from the regular-video condition, it may suggest that the infant paradigm is simply insensitive to infants' abilities to benefit from word-object relations. However, should infants benefit from regular word-object relations in Experiment 3, it will indicate important developmental differences in infants' abilities to integrate audio and visual information in a statistical learning task.

Method

Participants

Participants were 45 infants between the ages of 19.5 and 20.5 months ($M = 20.12$). Infants were randomly assigned to one of three groups: no-video, regular-video, and irregular-video. To obtain data from 45 infants, it was necessary to test 63. The additional 18 infants were excluded for the following reasons: fussing or crying (16), test trial looking times averaging less than 3 seconds (1), or parental interference (1). According to parental report, all infants were full term, and free of ear infections at the time of testing.

Stimuli

The stimuli were identical to those in Experiment 2.

Procedure

The procedure was identical to that of Experiment 2.

Results

Infants in the no-video condition looked at word trials for 8.3 sec ($SE = 0.6$), and at part-word trials for 8.2 sec ($SE = 0.5$). This difference in looking trials between words and part-words was not significant, $t(14) < 1$. Like the younger infants in Experiment 2, 20-month-olds in the no-video condition failed to distinguish between words and part-words, showing no evidence of learning. At neither age should this be taken as evidence that infants are unable to learn from audio stimuli alone – prior experiments clearly demonstrate that infants are able to do so (e.g., Saffran et al., 1996). Infants' failure in the current experiments is due to the fact that the stimuli are presented much more briefly than in prior experiments. While infants can learn from stimuli presented for this duration, they may only do so for natural – as opposed to synthesized – speech (e.g., Thiessen et al., 2005).

Infants in the irregular-video condition also showed no significant preference, looking equivalently long at word trials ($M = 9.6$, $SE = 0.7$) and part-word trials ($M = 9.1$, $SE = 0.6$), $t(14) < 1$. Interestingly, unlike the 8-month-olds in Experiment 2, 20-month-olds did not appear to learn from the irregular-video condition. Note that 8-month-olds' looking times were much longer to both kinds of test trials ($M = 11.1$ sec) than that of the 20-month-olds. This may indicate that the testing situation was more interesting to 8-month-olds than 20-month-olds. Sustained attention to the input facilitates statistical learning (e.g., Toro et al., 2005). The 20-month-olds in the current experiment may simply have failed to attend to the stimuli long enough to learn.

Regardless of infants' performance in the other two conditions, the question that motivated this experiment was whether they are facilitated in learning from the regular-video stimuli. Infants in the regular-video condition looked at word test trials for 7.2 sec ($SE = 0.6$), and to part-word test trials for 8.5 sec ($SE = 0.7$). This difference was significant, $t(14) = 2.3$, $p < .05$. Only infants in the regular-video condition showed evidence of learning; no other group demonstrated the ability to distinguish between words and part-words. A series of planned 2 x 2 ANOVAs comparing looking times across conditions assessed this more rigorously. In none of the ANOVAs was there a significant main effect of condition, nor of test trial (all Fs < 1). Similarly, there was no interaction between condition and test trial when comparing participants in the no-video condition to participants in the irregular-video condition ($F < 1$).

Most importantly, though, there were significant interactions between condition and looking time when comparing participants in the regular-video condition to participants in both the no-video condition [$F(1, 28) = 3.2$, $p < .05$] and the irregular-video condition [$F(1, 28) = 3.4$, $p < .05$]. These interactions indicate that infants' preference in the regular-video condition was significantly different from their lack of preference in either of the other two conditions. This confirms that 20-month-olds, like adults, performed significantly better in the regular-video condition than either of the other two conditions. For children at this age, complexity can facilitate learning by providing multiple learnable regularities in the input. This suggests an important developmental difference between 8- and 20-months of age, with only 20-month-olds showing the ability to benefit from regular audio-video relations in a manner comparable to adults.

General Discussion

One of the reasons that language is such an effective communicative tool is that it allows speakers to express multiple pieces of information simultaneously. For example, a simple observation about the state of the world, such as "the Pirates won," is coupled with affective information that indicates how the speaker feels about that state of affairs. This means that language is rich in possible relations for learners to discover, both between aspects of the speech signal (such as words and pitch), and between speech and meaning. Indeed, infants are able to detect many of these possible relations (e.g., Fisher, Klinger, & Song, 2006; Saffran et al., 1996; Smith & Yu, 2008; Thiessen & Saffran, 2007). However, the need for constraints is necessary for learners presented with rich input, especially input in which multiple relations are present simultaneously.

The current results indicate that at least one of those constraints is a developmental constraint. The ability to integrate simultaneous audio and visual information in a statistical learning task develops during the first two years of life. Whereas adults benefit from the presence of regular word-object associations, 8-month-old infants do not. Critically, 20-month-olds, like adults, benefit from the regular relations between words and visual objects available in these stimuli. One possibility for this developmental difference relates to older children's vastly greater experience with word-object correspondences in language. Ongoing work with non-linguistic stimuli will assess this possibility. Though the current data do not differentiate between domain-specific maturational accounts (e.g., Waxman & Booth, 2000) and accounts that implicate more general processes, both kinds of accounts share an important commonality. On either account, young infants are not learning as much as adults are when presented with stimuli in which words and objects co-occur. This may actually be beneficial for young learners. By preferentially detecting only some of the available relations in the stimuli, infant learners may avoid a combinatorial explosion (e.g., Newport, 1990).

References

Bahrick, L.E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology, 41*, 352-363.

Chambers, K.E., Onishi, K.H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experiences. *Cognition,* **87**, B69-B77.

Creel, S.C., Newport, E.L., & Aslin, R.N. (2004). Distant Melodies: Statistical learning of non-adjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 1119-1130.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.

Fenson, F., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (2002). *MacArthur Communicative Development Inventories: User's Guide and Technical Manual.* Baltimore: Paul H. Brookes.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181-195.

Fisher, C., Klingler, S.L., & Song, H. (2006). What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition, 101*, B19-B29.

Frick, J.E., & Richards, J.E. (2001). Individual differences in infants' recognition of briefly presented visual stimuli. *Infancy, 2*, 331-352.

Gold, M.E., (1967). Language identification in the limit. *Information and Control, 10*, 447-474.

Golinkoff, R.M., Shuff-Bailey, M., Olguin, R., & Ruan, W. (1995). Young children extend novel words at the basic level: Evidence for the principle of categorical scope. *Developmental Psychology, 31*, 494-507.

Graf Estes, K.M., 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.

Hall, G.D. (1991). Acquiring proper nouns for familiar and unfamiliar objects: Two-year-olds' word-learning biases. *Child Development, 62*, 1142-1154.

Hayes, J.R., & Clark, H.H. (1970). Experiments on the segmentation of an artificial speech analogue. In J. Hayes (Ed.) *Cognition and the Development of Language*, pp. 221-234. New York: John Wiley and Sons.

Hollich, G., Newman, R.S., & Jusczyk, P.W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development, 76*, 598-613.

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

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Landau, B., Smith, L.B., & Jones, S.S. (1988). The importance of shape in early lexical learning. *Developmental Psychology, 28*, 273-286.

Lewkowicz, D.J. (1986). Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior and Development, 9*, 335-353.

Lewkowicz, D.J. (2003). Learning and discrimination of audiovisual events in human infants: The hierarchical relation between intersensory synchrony and rhythmic pattern cues. *Developmental Psychology, 39*, 795-804.

Marcus, G.F., Vijayan, S., Bandi Rao, S., & Vishton, P.M. (1999). Rule learning in 7-month-old infants. *Science, 283*, 77-80.

Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*, 57-77.

Markman, E.M., & Wachtel, G.A. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121-157.

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*, 465-494.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex, 42*, 846-854.

Mintz, T.H. (2002). Category induction from distributional cues in an artificial language. *Memory and Cognition,* **30**, 678-686.

Neil, P.A., Chee-Ruiter, C., Scheier, C., Lewcowicz, D.J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science, 9*, 454-464.

Newport, E.L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Pelphrey, K.A., & Reznick, J.S. (2003). Working memory in infancy. In R. Kail (Ed.), *Advances in Child Development and Behavior, 31*, pp. 173-227.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: MIT Press.

Quine, W.V.O. (1964). *Word and Object*. Cambridge, MA: MIT Press.

Rohde, D.L., & Plaut, D.C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition, 72*, 67-109.

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

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 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*, 27-52.

Saffran, J.R., & Thiessen, E.D. (2003). Pattern induction by infant language learners. *Developmental Psychology, 39*, 484-494.

Saffran, J.R., Newport, E.L., Aslin, R.N., Tunick, R.A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*, 101-105.

Saffran, J.R., & Wilson, D.P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy, 4*, 273-284.

Sell, A., & Kaschak, M.P. (under review). Does speech reading affect word segmentation?

Smith, L.B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*, 1558-1568.

Soto-Faraco, S., Navarra, J., Weikum, W.M., Vouloumanos, A., Sebastian-Galles, N., & Werker, J.F. (2007). Discriminating languages by speech-reading. *Perception and Psychophysics, 69*, 218-231.

Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381-382.

Thiessen, E.D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language, 56*, 16-34.

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

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

Voulomanos, A. (2007). Using probabilities to build a lexicon. Paper presented at the Calgary Workshop on Current Issues in Language Acquisition: Artificial and Statistical Language Learning, June 2007.

Waxman, S.R., & Booth, A.E. (2000). Principles that are invoked in the acquisition of words, but not facts. *Cognition, 77*, B33-B43.

Werker, J.F., Cohen, L.B., Lloyd, V.L., Casasola, M., & Stager, C.L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289-1309.