PAPER

Isolated words enhance statistical language learning in infancy

Casey Lew-Williams,[1] Bruna Pelucchi[1,2] and Jenny R. Saffran[1]

1. *Department of Psychology and Waisman Center, University of Wisconsin-Madison, USA*
2. *Department of Evolutionary Biology, University of Ferrara, Italy*

Abstract

Infants are adept at tracking statistical regularities to identify word boundaries in pause-free speech. However, researchers have questioned the relevance of statistical learning mechanisms to language acquisition, since previous studies have used simplified artificial languages that ignore the variability of real language input. The experiments reported here embraced a key dimension of variability in infant-directed speech. English-learning infants (8–10 months) listened briefly to natural Italian speech that contained either fluent speech only or a combination of fluent speech and single-word utterances. Listening times revealed successful learning of the statistical properties of target words only when words appeared both in fluent speech and in isolation; brief exposure to fluent speech alone was not sufficient to facilitate detection of the words' statistical properties. This investigation suggests that statistical learning mechanisms actually benefit from variability in utterance length, and provides the first evidence that isolated words and longer utterances act in concert to support infant word segmentation.

Introduction

For 15 years, infant statistical learning has been a focus of research in the area of cognitive development. Many researchers have shown that infants have a prodigious and domain-general ability to track statistical patterns in their environments (e.g. Aslin, Saffran & Newport, 1998; Fiser & Aslin, 2002; Kirkham, Slemmer & Johnson, 2002; Pelucchi, Hay & Saffran, 2009; Saffran, Aslin & Newport, 1996). Given sequential events X and Y, infants are highly sensitive to the probability that event X will transition to event Y (and vice versa). This sensitivity to transitional probabilities (TPs) has been robustly demonstrated in several auditory and visual domains. For example, when exposed to continuous speech with no pause-defined cues to word boundaries, infants can segment words by noting which syllables occur together with the most consistency.

A majority of these studies have assessed learning using highly simplified artificial languages. While the materials used in statistical language studies typically afford a great deal of experimental control, they fail to capture many of the dimensions of variability that characterize infant-directed speech (though see Pelucchi *et al.*, 2009; Singh, Nestor & Bortfeld, 2008; Thiessen, Hill & Saffran, 2005). This has led several researchers to suggest that statistical learning mechanisms cannot handle the complexities of real language, and to question the hypothesis that statistical learning launches infants

into language acquisition (Johnson & Tyler, 2010; Yang, 2004). Here, we aimed to capture a significant source of variability in infant-directed speech – the presence of both isolated words and longer utterances – to assess the degree to which statistical learning abilities are well matched to the structure of early language input.

Natural speech is replete with pauses, which in some cases provide valuable cues to the locations of word boundaries. Corpus analyses show that while caregivers often produce long, uninterrupted strings of words, they also produce much shorter utterances, such as words in isolation. These single-word utterances, which by definition have pauses at their edges, are a salient feature of infant-directed speech. Brent and Siskind (2001) found that 9% of caregiver utterances consisted of isolated words, of which 27% occurred two or more times in neighboring utterances (e.g. *See the doggie over there? Doggie!*). Fernald and Morikawa (1993) found similar proportions: an average of 9% of utterances spoken by American and Japanese mothers were single content words. Other researchers have estimated similar or even higher proportions of isolated words (Aslin, Woodward, LaMendola & Bever, 1996; Siskind, 1996; van de Weijer, 1998). Researchers disagree about whether this is a substantial proportion of early language input. But viewed through a wide lens, the presence of isolated words seems more impressive: by combining estimates of total language exposure by Hart and Risley (1995) and of isolated word exposure by Brent and

Address for correspondence: Casey Lew-Williams, Waisman Center, Infant Learning Lab, 1500 Highland Ave., Room 501, University of Wisconsin-Madison, Madison, WI 53705-2280, USA; e-mail: lewwilliams@wisc.edu

Siskind (2001), we can estimate that a child may hear up to 4 million isolated word tokens by their fourth birthday.

Given these numbers, isolated words could be expected to play a role in early language acquisition. At the perceptual level, many studies have demonstrated that after hearing words in sentences, infants can recognize them when they appear in isolation (e.g. Jusczyk & Aslin, 1995; Jusczyk & Hohne, 1997; Pelucchi *et al.*, 2009). Conversely, after hearing words in isolation, infants can recognize them when they appear within sentences (e.g. Gout, Christophe & Morgan, 2004; Houston & Jusczyk, 2000; Jusczyk & Aslin, 1995). Of particular interest are findings demonstrating that caregivers' use of isolated words is linked to later vocabulary development. Brent and Siskind (2001) found that 75% of 18-month-olds' first words had previously appeared in isolation in mothers' speech, and moreover, that the frequency of hearing a word in isolation (but not the frequency of hearing a word overall) was a unique predictor of later word use (see also Ninio, 1992).

Given the presence of isolated words in infant-directed speech, what is their role in statistical learning? To date, we do not know whether these acoustically demarcated units have any effect on infants' computations of statistically coherent syllable sequences, or more broadly speaking, whether statistical learning mechanisms can help infants make sense of natural speech that contains a mix of isolated words and longer utterances. Do isolated words offer a shortcut for word segmentation, such that infants do not need to track TPs in natural speech? Alternatively, perhaps the presence of isolated words facilitates the detection of TPs in fluent speech.

To address this issue, we tested 8- to 10-month-old English-learning infants' abilities to segment words from natural Italian speech that either did or did not contain isolated words. In Experiment 1, infants were tested in one of two conditions. In the *Fluent Speech* condition, infants heard a set of Italian sentences abbreviated from the materials used in a previous study of statistical learning (Pelucchi *et al.*, 2009, Experiment 3). Two types of target words were embedded within the sentences: high-TP words (TP = 1.0), where the component syllables never appeared elsewhere in the corpus, and low-TP words (TP = 0.33), where the component syllables recurred within many other words in the corpus. This brief, pause-free familiarization was designed to minimize the likelihood that infants would be able to learn enough about high- vs. low-TP words. In the *Fluent Speech + Isolated Words* condition, infants heard the same target words within sentences, and also heard the target words spoken in isolation between sentences. Critically, this condition was designed such that successful learning of the words' statistical properties required infants to integrate their experience hearing fluent speech and isolated words.

After familiarization with the Italian speech, infants were tested on their listening times to high- and low-TP words using the Headturn Preference Procedure. If infants cannot detect the words' statistical coherence based on a very short exposure to fluent speech alone, then we would expect listening times to be equivalent for both types of target words in the *Fluent Speech* condition. In the *Fluent Speech + Isolated Words* condition, the same results would emerge if infants relied on isolated words alone, since all target words appeared in isolation with the same frequency. However, we predicted an interaction between condition and the type of target word. If infants can integrate statistical information gained from hearing words in isolation and in fluent speech, then we would expect listening times to differ for high- and low-TP words in the *Fluent Speech + Isolated Words* condition.

Experiment 1

Method

Participants

Participants were 40 healthy, full-term infants with a mean age of 9.5 months (range = 8.3–10.9). Infants were from monolingual English-speaking households, and had no prior exposure to Italian or Spanish. Parents reported no pervasive developmental delays and no history of hearing or visual problems. Infants were randomly assigned to the *Fluent Speech* condition (10 female, 10 male) or the *Fluent Speech + Isolated Words* condition (10 female, 10 male). Eleven additional infants were tested but not included in analyses due to fussiness (10) or inattentiveness (1).

Speech stimuli

Infants were familiarized with natural, infant-directed Italian speech that either did or did not contain isolated tokens of target words (see Figure 1). In the *Fluent Speech* condition, infants heard grammatically correct and semantically meaningful sentences adapted from a previous study of statistical learning (Pelucchi *et al.*, 2009, Experiment 3). Sentences were recorded by a female native speaker of Italian at a rate of approximately six syllables per second, presented at approximately 60 dB_{SPL}. Embedded within the sentences were four bisyllabic target words: *casa*, *bici*, *fuga*, and *melo* (English translations: house, bike, escape, and apple tree). The target words appeared with equal frequency during familiarization, but their internal TPs differed. For half of the infants in this condition, *casa* and *bici* were high-TP words: their component syllables (*ca*, *sa*, *bi*, and *ci*) always appeared in the contexts of the words *casa* and *bici*. Thus the internal TPs of these words were 1.0. For these infants, *fuga* and *melo* were low-TP words: the initial syllables (*fu* and *me*) appeared 33% of the time in the words *fuga* and *melo*, but 67% of the time else-

Fluent Speech		Isolated Words	
Torno a **casa** con le **bici ca**riche di frutta in **bi**lico sulla sella.		**Casa**.	
Per ascoltare la *fuga* quasi **ca**di sul *melo* e inciampi sulla **bi**ro sull'erba.		*Fuga*.	
Il picchio si a**bi**tua a fare la sua **casa** in ogni *melo* **ca**vo e alto.		*Melo*.	
Di rado una **bici** in rapida *fuga* rincorre la moto **bi**gia e rossa.	300-ms pause	**Bici**.	300-ms pause
La **ca**via **Bi**da è in *fuga* da **casa** per aver gio**ca**to con le **bi**lie blu.		**Casa**.	
La **bici** ha su**bi**to un danno dentro la **casa** del **ca**po di Lara.		**Bici**.	
Se porti il *melo* sulla **bici** forse **ca**li un po' di chili.		*Melo*.	
Il tuo *melo* ar**ca**no *fuga* l'afa che de**bi**lita la folla.		*Fuga*.	
Gusto i **bi**goli dentro **casa** o cori**ca**ta all'ombra del *melo* verde.		*Melo*.	
Arriviamo in **bici** fino al **bi**vio del grande *melo* con un **ca**ro amico.		**Bici**.	
La zia **Ca**rola si è esi**bi**ta in una *fuga* colla **bici** verde.		*Fuga*.	
La **bi**scia in lenta *fuga* dal giardino **ca**pita in **casa** mia.		**Casa**.	

Figure 1 *Subset of Italian stimuli from Experiment 1. In the Fluent Speech + Isolated Words condition, participants heard both fluent speech and isolated words (left and right columns); in the Fluent Speech condition, participants heard only fluent speech (two repetitions of the sentences in the left column). High-TP words are italicized; low-TP words are bolded, as are additional occurrences of the initial syllables of low-TP words. Only one of the two familiarization corpora is shown in the figure. Filler sentences are omitted.*

where in the speech (always in stressed position). Thus the internal TPs of *fuga* and *melo* were 0.33. The other half of the infants heard a different, counterbalanced familiarization, with *casa* and *bici* as low-TP words and *fuga* and *melo* as high-TP words.

In Pelucchi *et al.* (2009), infants successfully discriminated high-TP from low-TP words after receiving 18 exposures to each target word. Importantly, in the present experiment, infants received just 12 exposures to each target word. This reduction was designed to decrease the likelihood that infants would successfully segment target words when hearing only fluent speech.

In the *Fluent Speech + Isolated Words* condition, infants received only six exposures to each target word within sentences, with 12 additional occurrences of the initial syllables of low-TP words. Each target word was also presented three times in isolation, yielding a total of nine exposures to each target word. Additional filler sentences were interspersed in order to match familiarization length across conditions. In order to mimic caregiver use of isolated words (Brent & Siskind, 2001), each isolated word appeared between sentences, bordered by 300-ms pauses, and always followed a sentence that contained the same target word. Test items for both conditions consisted of the four target words (two high-TP, two low-TP); the tokens were identical to the single-word utterances used in the *Fluent Speech + Isolated Words* condition.

Procedure

The experiment consisted of a familiarization phase (90 sec) and a test phase (approximately 3 min), using the Headturn Preference Procedure. Infants were seated on a caregiver's lap within a double-walled, sound-attenuated booth, equipped with a center light and two side-lights. Infants heard speech stimuli while parents listened to unrelated speech over headphones. The experimenter coded infant head-turns outside of the booth via a closed-circuit camera using a button-box connected to a computer. During familiarization, infants heard Italian speech from laterally placed speakers. The lights flashed contingently on infants' looking behavior. Each test trial was initiated when infants looked to a flashing center light, at which point the experimenter signaled the center light to extinguish and one of the side-lights to begin flashing. When infants looked at the side-light, an isolated token of a high- or low-TP word was repeated until infants looked away for 2 s or 20 s had elapsed. Side of stimulus presentation was counterbalanced, and the experimenter was blind to test-item presentation. Infants' looking time was calculated automatically using custom-designed software. During the test, each high- and low-TP word was repeated on three trials, randomized by block, for a total of 12 test trials.

Results and discussion

Looking times were analyzed in a 2 × 2 mixed ANOVA with condition (*Fluent Speech*, *Fluent Speech + Isolated Words*) as a between-subjects factor, and test item (high-TP, low-TP) as a within-subjects factor (see Figure 2).[1] Looking times revealed a significant main effect of test item, $F(1, 38) = 4.3$, $p < .05$, $\eta_p^2 = .10$, suggesting that infants listened significantly longer to high-TP words than to low-TP words. Analyses also showed a significant interaction between condition and test item, $F(1, 38) = 7.8$, $p < .01$, $\eta_p^2 = .17$. Bonferroni-corrected post-hoc comparisons ($p < .025$) revealed that infants in the *Fluent Speech* condition listened equally to high- and low-TP words, $t(19) = .5$, $p = .6$, $d = .09$, but that infants in the *Fluent Speech + Isolated Words* condition listened significantly longer to high-TP words, $t(19) = 3.2$, $p < .005$, $d = .53$. Given this very brief exposure, infants

---

[1] In both Experiments 1 and 2, performance did not vary by gender and did not differ across the two counterbalanced corpora. Thus, these factors were not included in the main analyses.
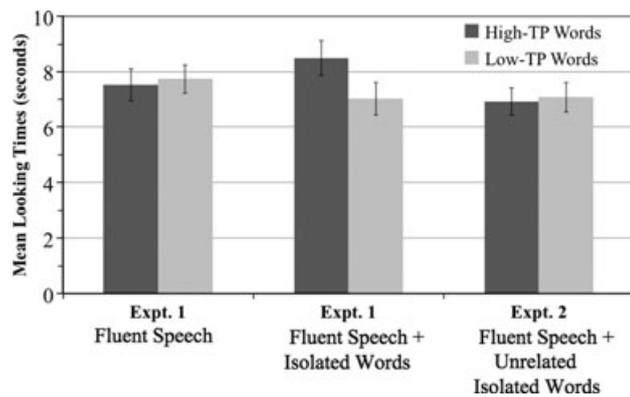
Figure 2 *Mean looking times to high-TP and low-TP words in Experiment 1 (Fluent Speech condition, Fluent Speech + Isolated Words condition) and Experiment 2 (Fluent Speech + Unrelated Isolated Words condition). Error bars represent standard errors of the mean.*

failed to learn enough about the words' statistical properties to successfully discriminate the test items in the absence of isolated words. However, the presence of single-word utterances facilitated infants' detection of the statistical properties of those words in fluent speech.

Infants' successful discrimination of high- and low-TP words in the *Fluent Speech + Isolated Words* condition is particularly striking given the frequency with which infants were exposed to target words across conditions. Infants in this condition received just nine exposures to each target word, while infants in the *Fluent Speech* condition received 12 exposures. Even with fewer opportunities to track statistical properties of Italian speech, infants integrated information gained from hearing words in isolation with statistical information available in fluent speech. Critically, infants heard high- and low-TP words equally often in isolation. If they had attended only to isolated words, we would expect no difference in looking to each word type, since all target words would have an identical TP of 1.0. Thus, attention to the statistical coherence of syllables was required for infants to successfully discriminate high- from low-TP words.

An alternative explanation for these findings is that the presence of isolated words simply made the speech more engaging. We know that infants prefer listening to speech with variability in pitch (Cooper & Aslin, 1990; Fernald & Kuhl, 1987), and that pitch variability heightens infants' ability to track sequential statistics (Thiessen *et al.*, 2005). It is thus possible that variability in utterance length heightened infants' interest in listening to the Italian familiarization. This may have resulted in an increased ability to detect statistical properties of target words from fluent speech alone, regardless of the contribution of isolated words. To test this alternative hypothesis, infants in Experiment 2 listened to a combination of fluent Italian speech and isolated words, but there was no phonological overlap between the words presented in fluent speech and isolation. If variability alone enhances performance, then infants should con-

tinue to show successful discrimination in Experiment 2. However, if the crucial factor is the phonological overlap between the isolated words and the fluent speech, then infants should fail to discriminate between the test items in Experiment 2.

Experiment 2

Method

Participants

Participants were 20 healthy, full-term infants (10 female) with a mean age of 9.4 months (range = 8.4–10.8). Six additional infants were tested but not included in analyses due to fussiness (4) or inattentiveness (2).

Speech stimuli and procedure

Materials and procedures were comparable to the *Fluent Speech + Isolated Words* condition from Experiment 1. In both experiments, there were nine exposures to the target words *casa*, *bici*, *fuga*, and *melo* during the familiarization (90 s). In the previous experiment, infants heard each target word six times in fluent speech and three times in isolation; however, in Experiment 2, all nine exposures were in fluent speech, and infants heard *different* Italian words in isolation: *pane*, *dono*, *tema*, and *socio*. Like the target words, these unrelated isolated words were bisyllabic and conformed to a strong–weak stress pattern. Crucially, however, these isolated words did not overlap with the target words.

Results and discussion

Listening times in Experiment 2 were compared to listening times in the *Fluent Speech + Isolated Words* condition from Experiment 1 in a 2 (condition) × 2 (test item) mixed ANOVA. Analyses revealed a significant interaction, $F(1, 38) = 5.6$, $p < .025$, $\eta_p^2 = .13$, suggesting that listening times to high- and low-TP words differed in the two conditions. A paired-samples *t*-test showed that infants did not show a significant difference in listening times to high- and low-TP words in Experiment 2, $t(19) = .3$, $p = .8$, $d = .07$ (see Figure 2). The lack of discrimination suggests that isolated words do not simply enhance attention to the speech stream. Rather, caregiver use of isolated words seems to play a role in helping infants detect statistical regularities related to those words in otherwise pause-free speech.

General discussion

The usefulness of statistical learning mechanisms in explaining early language learning is dependent upon infants' abilities to organize the variability inherent in

real speech. We focused here on a particularly salient source of variability in caregiver speech – the presence of both isolated words and longer utterances – and provided the first evidence that isolated words and longer utterances act in concert to facilitate infant word segmentation. In two preferential-listening experiments, English-learning infants received brief exposure to Italian. When hearing high- and low-TP target words in a brief corpus of fluent speech alone, infants did not demonstrate that they learned enough about the words' statistical properties to distinguish between them. However, infants who heard a similarly brief corpus containing target words both in fluent speech and in isolation successfully detected the coherence of their component syllables. This finding indicates that statistical learning mechanisms may in fact benefit from variability in utterance length. The only way that infants could have discriminated high- from low-TP words in this condition was through the interaction of fluent speech and isolated words.

The industry of studies on statistical language learning was built on the premise that pauses do not reliably occur between words, posing a difficult puzzle of word segmentation to young language learners. Previous research shows that infants are highly adept at discovering statistical regularities in pause-free speech (Saffran *et al.*, 1996), and that embedding words in sentences has many other benefits in early language learning, by helping infants recognize words faster (Fernald & Hurtado, 2006), anticipate subsequent words (Lew-Williams & Fernald, 2007), discover grammatical categories (Mintz, 2003), and produce multiword speech (Bannard & Matthews, 2008).

Isolated words – which by definition are acoustically demarcated from other utterances by moments of silence – may also have positive consequences for language learning (Peters, 1983; Pinker, 1984). Using computational or adult experimental data, previous researchers have demonstrated that isolated words facilitate the discovery of word boundaries in fluent speech (Brent & Cartwright, 1996; Conway, Bauerschmidt, Huang & Pisoni, 2009; Dahan & Brent, 1999; van de Weijer, 1998). Descriptions of this process converge on a common premise: words spoken in isolation become perceptually familiar, 'pop out' when heard in fluent speech, and provide a means of segmenting adjacent words in fluent speech. Similarly, infant studies suggest that highly familiar words, such as the infant's name or the family's word for mother, provide useful cues to the boundaries of the words adjacent to them (Bortfeld, Morgan, Golinkoff & Rathbun, 2005), and even help infants in challenging artificial language tasks (Mersad & Nazzi, 2010). However, this hypothesis has not yet been tested using actual isolated words with infant learners.

The experiments reported here offer a new perspective on how a combination of isolated words and longer utterances may support early word segmentation. Words heard in isolation are relatively familiar and recognizable when heard in fluent speech, and they enhance infants' abilities to detect the statistical properties of syllable sequences in the global language input. Importantly, isolated words are not likely to be essential for language learning, and they do not simply replace the need for infants to compute the locations of word boundaries in fluent speech. But isolated words are a prevalent feature of caregiver speech, and even a minimal number of exposures to words in isolation may confer advantages in early statistical learning. This research thus complements and extends previous work on the importance of infant-directed speech (Brent & Siskind, 2001; Gleitman, Newport & Gleitman, 1984; Kemler-Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989; Singh, Nestor, Parikh & Yull, 2009; Thiessen *et al.*, 2005; Werker, Pons, Dietrich, Kajikawa, Fais & Amano, 2007).

Isolated words appear to facilitate learning about statistical attributes of fluent speech. However, both long sentences and isolated words are very much part of the statistical landscape of speech, and their relationship is likely to be symbiotic: the distributions of syllables in fluent speech could also help infants learn about the isolated words themselves. While infants may treat a novel isolated word like *casa* as a single, unsegmented unit, they may also oversegment, treating *casa* as two monosyllabic words. Recent computational models of word segmentation have focused on this issue, examining the initial assumptions that learners could make about the nature of words (Goldwater, Griffiths & Johnson, 2009). Generally, ideal learners tend to favor the most efficient segmentation strategies given the input data, which in this case would be treating *casa* as a single word. But statistics are needed to guide learners toward this correct interpretation. Over time, infants could gradually determine that the component syllables *ca* and *sa* should be more tightly linked by noting the regularity with which those syllables co-occur in the language environment. Thus, while isolated words support statistical learning in long utterances, statistical features of long utterances also help infants learn about the structure of isolated words.

While isolated words are particularly salient in the speech stream, they are just one in a network of useful segmentation cues that may inform each other multidirectionally. It will be important to discover how isolated words work together with other linguistic and domain-general cues to signal the locations of word boundaries, such as lexical stress (Curtin, Mintz & Christiansen, 2005; Johnson & Jusczyk, 2001), phrase-level prosody (Shukla, Nespor & Mehler, 2007), sentence position (Seidl & Johnson, 2006), language-specific phonotactic constraints (Brent & Cartwright, 1996), and concurrently presented visual information (Cunillera, Cámara, Laine & Rodríguez-Fornells, 2010). Future studies should also aim to understand how statistical learning mechanisms operate given other dimensions of variability in infant-directed speech, such as the presence of multiple talkers in complex turn-taking interactions. In some cases, as in

the research on utterance length reported here, variability does not pose a problem to the young learner, but in other cases, variability may disrupt learning. Studies of the variability inherent in infant-directed speech will both enhance and constrain accounts of early language acquisition, by allowing us to determine the degree to which infant learning mechanisms are well matched to the structure of early language input.

Acknowledgements

References

Aslin, R.N., Saffran, J.R., & Newport, E.L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, **9**, 321–324.

Aslin, R.N., Woodward, J.Z., LaMendola, N.P., & Bever, T.G. (1996). Models of word segmentation in fluent maternal speech to infants. In K. Demuth & J.L. Morgan (Eds.), *Signal to syntax* (pp. 117–134). Mahwah, NJ: Erlbaum.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: the effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, **19**, 241–248.

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

Brent, M.R., & Cartwright, T.A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, **61**, 93–125.

Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, **81**, B33–B44.

Conway, C.M., Bauerschmidt, A., Huang, S.S., & Pisoni, D.B. (2009). Implicit statistical learning in language processing: word predictability is the key. *Cognition*, **114**, 356–371.

Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech within the first month after birth. *Child Development*, **61**, 1584–1595.

Cunillera, T., Cámara, E., Laine, M., & Rodríguez-Fornells, A. (2010). Speech segmentation is facilitated by visual cues. *Quarterly Journal of Experimental Psychology*, **63**, 260–274.

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

Dahan, D., & Brent, M.R. (1999). On the discovery of novel wordlike units from utterances: an artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*, **128**, 165–185.

Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, **9**, F33–F40.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, **10**, 279–293.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, **64**, 637–656.

Fiser, J., & Aslin, R.N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **28**, 458–467.

Gleitman, L., Newport, M., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, **11**, 43–79.

Goldwater, S., Griffiths, T.L., & Johnson, M. (2009). A Bayesian framework for word segmentation: exploring the effects of context. *Cognition*, **112**, 21–54.

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.

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

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

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., & Tyler, M.D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, **13**, 339–345.

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

Jusczyk, P.W., & Hohne, E.A. (1997). Infants' memory for spoken words. *Science*, **277**, 1984–1986.

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

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, **83**, B35–B42.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, **18**, 193–198.

Mersad, K., & Nazzi, T. (2010). When Mommy comes to the rescue of statistics. Poster presented at the Biennial Meeting of the International Society for Infant Studies, Baltimore, MD.

Mintz, T.H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, **90**, 91–117.

Ninio, A. (1992). The relation of children's single word utterances to single word utterances in the input. *Journal of Child Language*, **19**, 87–110.

Pelucchi, B., Hay, J.F., & Saffran, J.R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, **80**, 674–685.

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

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

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

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

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, **54**, 1–32.

Singh, L., Nestor, S.S., & Bortfeld, H. (2008). Overcoming the effects of variation in infant speech segmentation: influences of word familiarity. *Infancy*, **13**, 57–74.

Singh, L., Nestor, S.S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, **14**, 654–666.

Siskind, J.M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, **61**, 39–91.

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

van de Weijer, J. (1998). *Language input for word discovery*. Nijmegen: Max Planck Institute for Psycholinguistics.

Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, **103**, 147–162.

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