HHS Public Access
Author manuscript
J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2019 February 01.

Published in final edited form as:
J Exp Psychol Learn Mem Cogn. 2018 February ; 44(2): 221–232. doi:10.1037/xlm0000448.

THE LONGEVITY OF STATISTICAL LEARNING: WHEN INFANT MEMORY DECAYS, ISOLATED WORDS COME TO THE RESCUE

Ferhat Karaman and **Jessica F. Hay**
Department of Psychology, University of Tennessee, Knoxville, TN

Abstract

Research over the past two decades has demonstrated that infants are equipped with remarkable computational abilities that allow them to find words in continuous speech. Infants can encode information about the transitional probability (TP) between syllables to segment words from artificial and natural languages. As previous research has tested infants immediately after familiarization, infants' ability to retain sequential statistics beyond the immediate familiarization context remains unknown. Here, we examine infants' memory for statistically defined words 10 minutes following familiarization with an Italian corpus. Eight-month-old English-learning infants were familiarized with Italian sentences that contained four embedded target words – two words had high internal TP (HTP, TP=1.0) and two had low TP (LTP, TP=.33) – and were tested on their ability to discriminate HTP from LTP words using the Headturn Preference Procedure. When tested following a 10-minute delay, infants failed to discriminate HTP from LTP words, suggesting that memory for statistical information likely decays over even short delays (Experiment 1). Experiments 2–4 were designed to test whether experience with isolated words selectively reinforces memory for statistically defined (i.e., HTP) words. When 8-month-olds were given additional experience with isolated tokens of *both* HTP and LTP words immediately after familiarization, they looked significantly longer on HTP than LTP test trials 10 minutes later. Although initial representations of statistically defined words may be fragile, our results suggest that experience with isolated words may reinforce the output of statistical learning by helping infants create more robust memories for words with strong versus weak co-occurrence statistics.

Keywords

statistical learning; memory; word segmentation; isolated words

---

To become proficient in their native language, infants must be able to discover linguistic patterns in input that lacks explicitly accessible structure. For example, words – often described as the building blocks of language – are typically produced in the context of a continuous speech stream. Unlike written language, spoken language does not contain silences that reliably demarcate word boundaries (Cole & Jakimik, 1980), yet infants begin recognizing words within their first year of life (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012). By 9 months of age, infants show sensitive to a variety of language-

Corresponding Author. Jessica F. Hay, 416D Austin Peay Building, The University of Tennessee, Knoxville, TN 37996-0900 USA, (865) 974-3329, jhay@utk.edu.

specific cues to word boundaries including phonotactic regularities (e.g., Friederici & Wessels, 1993; Mattys & Jusczyk, 2001), allophonic variation (e.g., Jusczyk, Hohne, & Bauman, 1999) and native language prosodic patterns (e.g., English: Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Houston, & Newsome, 1999; Dutch: Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000; German: Höhle & Weissenborn, 2003; French: Goyet, Nishibayashi, & Nazzi, 2013; Spanish/Catalan: Bosch, Figueras, Teixidó, & Ramon-Casas, 2013). Even before infants show sensitivity to language-specific cues, they are remarkably good at making use of language-general word boundary cues (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996). Specifically, by around 6.5 to 8 months of age infants appear to be able to find word-like units in a continuous stream of speech by tracking the sequential statistics, namely the transitional probability (TP, the probability of X given Y in the sequence XY), between syllables in both artificial (Aslin et al., 1998; Saffran et al., 1996; Thiessen & Saffran, 2003) and natural languages (Pelucchi, Hay, & Saffran, 2009a, 2009b). The ability to track TP in fluent speech has been observed in infants from various language backgrounds including English (e.g., Aslin et al., 1998; Hay & Saffran, 2012; Saffran et al., 1996), French (e.g., Goyet, Nishibayashi, & Nazzi, 2013; Mersad & Nazzi, 2012) and Dutch (Johnson & Tyler, 2010), to name a few. Once discovered, these word-like units function as candidate labels for objects in word learning tasks (Graf Estes, Evans, Alibaba, & Saffran, 2007; Hay, Pelucchi, Graf Estes, & Saffran, 2011). Thus, infants' ability to track statistical patterns in their input may be a partial explanation for how infants learn language so quickly. These sophisticated computational abilities may be of little help building a vocabulary, however, unless the sound patterns of recently extracted words are encoded into long-term memory and remembered over time and in a variety of different contexts.

While it is now generally accepted that memory is a fundamental cognitive capacity that begins developing early in infancy (Rovee & Fagen, 1976; Rovee-Collier, 1999), prior to the 1970s researchers did not believe that infants had the capacity to form stable representations of information. With the development of sophisticated non-verbal tasks (e.g., the high-amplitude sucking procedure, the mobile-kicking paradigm, the deferred imitation paradigm, and the visual preference task) researchers, starting with the work of Rovee and Fagen (1976; see also Rovee-Collier, 1999) and Meltzoff (1985, 1988), have successfully demonstrated that infants not only have the ability to form memories, but also to remember relevant events, actions, and information over a substantial period of time. In addition to behavioral research, studies from behavioral neuroscience (using animal models) (e.g., Nakashiba et al., 2008; Squire, 1992) and developmental neuroscience (using neuroimaging and electrophysiology) (e.g., Nyberg & Cabeza, 2000) have given us insights into the ways in which memory, and the brain structures that support it, change across development (see Bauer, 2006; Gómez, 2017 for reviews). This work has revealed that memory is not a single simple faculty but consists of different systems (such as working memory/short-term memory, long-term memory) and sub-systems (such as explicit memory and implicit memory). Further, each type of memory involves a collection of various sub-processes (such as encoding, consolidation, storage, and retrieval). These memory processes are especially important during language learning and use. Thus, incorporating questions about memory in the study of early language acquisition is an important endeavor.

Even though early memory mechanisms undoubtedly play a vital role in supporting word learning (see Wojcik, 2013 for a recent review), only a handful of studies have actually examined whether infants remember words over time. In one study, Jusczyk and Hohne (1997) explored 8-month-old infants' long-term memory for the sound patterns in words by presenting infants with target words embedded in passages for 30 minutes a day for 10 days. When tested following a 2-week delay, infants preferred listening to the target words compared to foils, suggesting that they had successfully stored the sound patterns of these frequently presented words in their long-term memory. In a subsequent study, Houston and Jusczyk (2003) found that 7.5-month-old infants could remember the phonological patterns of words presented briefly in isolation (i.e., *cup* & *dog* or *feet* & *bike* 30 times each) when tested 24-hours later on passages produced by the same talker. Together, these studies suggest that repeated exposure to the phonological properties of words might be one factor that drives long-term memory.

There is also evidence that infants can remember and generalize grammatical patterns when tested following both shorter (e.g., 5 minute) and longer (e.g., 4 to 24 hour) delays. For example, Gómez and Gerken (1999) familiarized 12-month-old infants with a miniature artificial language and tested them on their ability to discriminate novel grammatical strings from illegal strings following a 5-minute delay. Across four experiments infants demonstrated successful learning of the grammatical patterns. Importantly, infants were able to retain these complex grammatical patterns over a 5-minute delay following only a brief 2-minute exposure to the artificial grammar. Further, 15-month-olds familiarized with an artificial language containing non-adjacent dependencies (e.g., A×B) appear to remember the non-adjacent dependencies when tested 4 hours later (Gómez, Bootzin, & Nadel, 2006). Interestingly, Gómez and colleagues (2006) also found that napping between familiarization and test promotes abstraction of these statistical patterns, whereby infants demonstrate generalization to similar, but not identical stimuli. In a follow up study, Hupbach, Gómez, Bootzin, and Nadel (2009) found that only infants who napped within 4 hours of familiarization were able to retain abstract information for 24 hours.

Despite more than twenty years of research on the role of infant statistical learning in speech segmentation, relatively little is known about infants' memory for statistically defined words, let alone which factors facilitate or impede memory for speech. Since infant statistical learning is typically assessed immediately after familiarization with a speech stream, we do not know whether memories for statistical patterns persist and how they affect later learning experiences. This necessarily limits our understanding of the extent to which infants derive benefit from their experience with the sequential statistics available in fluent speech. If syllables that tend to co-occur have a privileged status when infants next encounter them, then subsequent language learning might be facilitated.

In the present study, we assess infants' long-term memory for statistically defined words following familiarization with a naturally produced Italian corpus similar to ones used successfully in a number of other studies on statistical learning (see Hay et al., 2011; Lew-Williams, Pelucchi, & Saffran, 2011; Pelucchi et al., 2009a; 2009b). Here, and in these previous studies, infants were presented with a series of Italian sentences in which four target words (i.e., fuga, melo, casa, & bici) were embedded 18 times each. Two of the words had high TP (HTP; TP = 1.0) in that the first syllable always co-occurred with the second syllable; the syllables in these words did not occur anywhere else in the corpus. Two of the words had low TP (LTP; TP = 0.33) because the first syllable occurred in other words throughout the corpus. Italian was selected because it is sufficiently unfamiliar to infants, yet it shares a number of phonological and prosodic features with the infants' native language (i.e., English).

One thing to note is that in the natural Italian corpus used here, and in previous studies, HTP and LTP words occupy very different positions in the statistical landscape. The TPs of the HTP words are relatively high in the corpus (TP = 1.0) relative to the TPs in the surrounding landscape (TP ~ .25), creating local peaks/TP maxima. Thus, the internal TPs of the HTP words are much higher than the TPs that define the HTP word boundaries. On the other hand, the TPs of the LTP words (TP = .33) are not that different from the TPs in the surrounding statistical landscape. LTP words may be more difficult to segment from the corpus, both because of their low internal TP and because their internal TP is not much higher than the TPs that define the LTP word boundaries. Infants may find words in speech by attending to either high TP sequences, or by noticing when there is a dip in TP at word boundaries. Of course, infants may also be able to integrate these two sources of information. Either way, successful discrimination of HTP and LTP words indicates that infants have tracked co-occurrence patterns in the corpus.

In these original studies, 8-month-old infants were familiarized with approximately 2 minutes of a natural Italian corpus and were immediately tested using the Headturn Preference Procedure (Saffran et al., 1996) on their ability to differentiate HTP from LTP words (Pelucchi et al., 2009a, 2009b). In studies that have used artificial language materials (e.g., Saffran et al., 1996; Aslin et al., 1998), where infants are presented with 45–90 tokens of the target words, infants typically show a novelty preference, reflecting that they have discovered the words in the speech stream and are no longer interested in listening to them. In Pelucchi et al.'s first study (2009a), and in much of their subsequent work using natural Italian stimuli, infants showed a significant familiarity preference for HTP words over LTP words. Here, a familiarity preference suggests that infants have tracked the statistical regularities in the corpus and *remain* interested in listening to words that they recognize from the corpus. Familiarity preferences have been commonly observed in segmentation studies that have used natural language stimuli, beginning with Juscyzk and Aslin (1995). Although the direction of preference observed in infant research can vary based on a number of factors including the age of the infant, the amount of experience with and complexity of the stimuli, and the degree of processing (Hunter & Ames, 1988), it is generally accepted that any systematic group listening difference in these statistical learning studies reflects sensitivity to the statistical structure of the familiarization language. The findings that infants can track TP to discover words in these types of natural Italian corpora, have been replicated a number of times over the past few years (Hay et al., 2011; Lew-Williams et al., 2011; Pelucchi et al., 2009b).

Despite these important findings, little is known about infant memory for statistically defined sound sequences. Pulling words out of the speech stream is just the first step in learning a language. Building a vocabulary requires infants to remember the sound patterns of words heard in continuous speech, so that they can access those words during future language learning opportunities. Being able to remember words is particularly important as objects and concepts may not be in the infant's immediate environment when the sounds patterns of words are first encountered. This is especially true given that a large portion of early language input comes from overheard speech (Akhtar, 2005; Akhtar, Jipson, & Callanan, 2001). Demonstrating that infants are able to encode sequential statistics into memory and remember them over time will support theories of statistical learning as a process by which infants acquire language. Thus, in the current experiments we examine the longevity of statistical learning by testing infants' ability to encode the sound patterns of words extracted from continuous speech and remember them over time.

Experiment 1

In Experiment 1, twenty-four 8-month-old infants were first familiarized with a naturally produced Italian corpus used by Pelucchi and colleagues (2009a) and tested on their ability to discriminate words with high versus low TP following a 10-minute delay. We selected a 10-minute delay because it falls within a retention interval that would be relevant to infants' everyday experiences with hearing speech and seeing referents in the environment, and yet is a sufficiently long delay to ensure that performance would be based on retrieval from long-term memory. Although 10 minutes remains a relatively brief period of time, it far exceeds the limits of auditory short-term memory infancy (Ross-Sheehy & Newman, 2015). Further, since 8-month-old infants lack the relevant verbal and cognitive skills, it is highly unlikely that they are capable of deliberate rehearsal or other ways of keeping the stimulus active in working memory for 10 minutes (Ross-Sheehy, Oakes, & Luck, 2003; Cowan et al., 2005). Thus, infants will only be able differentiate HTP from LTP words following a 10-minute delay if the HTP words are encoded into, and retrieved from, long-term memory. To be able to compare our findings to those of previous work testing statistical learning in natural language (e.g., Pelucchi et al., 2009a, 2009b), we also tested 8-month-old infants. Further, 8-month-olds must actively engage in speech segmentation due to their limited receptive vocabulary. Thus, 8 months is a relevant age to examine memory for statistical regularities. Following the findings of Pelucchi and colleagues (Pelucchi et al., 2009a, 2009b), using similar familiarization materials, a familiarity preference was our index of learning.

Methods

Participants—Twenty-four healthy, full-term infants (8 males and 16 females) with a mean age of 8.4 months (range = 8.1 to 8.9) participated in Experiment 1. In Experiment 1, and in all subsequent experiments, infants were from monolingual English-speaking families with no prior exposure to Italian or Spanish, and had no history of hearing or vision impairments. Participants were recruited through the Child Development Research Group database maintained in the Department of Psychology at the University of Tennessee, Knoxville, and through community outreach initiatives in the greater Knoxville area. Twelve additional infants were tested but not included in the analyses for the following reasons: fussiness, including whimpering and/or continuous crying leading to a failure to complete at least 8 of the 12 test trials (n = 11) or not paying attention as reflected by failure to orient to the monitors during testing (n = 1). The Internal Review Board at the University of

Tennessee approved all recruitment procedures and experimental protocols used in Experiments 1–4. Parental consent was obtained for all participants. Infants received a small gift for their participation.

Stimuli—Speech stimuli were identical to those used successfully in a previous study of word segmentation (Pelucchi et al., 2009a; Experiment 3). The familiarization corpus consisted of 12 grammatically correct and semantically meaningful Italian sentences produced with a lively prosody by a female native speaker of Italian (see the Appendix for sentence lists). All sentences were matched in intensity to be presented at approximately 65 dB_{SPL}. During familiarization, infants heard three repetitions of the corpus for a total duration of 2 minutes 15 seconds. Two counterbalanced languages were created to control for any arbitrary listening biases at test.

Four disyllabic target words (*bici*, *casa*, *fuga*, and *melo*) were embedded in the speech stream. These target words were phonetically and phonotactically legal in English and all followed a strong/weak (trochaic) stress pattern. The target words appeared with equal frequency, occurring six times in each corpus, but their internal TPs differed. In Language A, the syllables *fu*, *ga*, *me*, and *lo*, appeared only in the words *fuga* and *melo*. Therefore, the internal TPs of *fuga* and *melo* were 1.0 (HTP words). However, the internal TPs of the other two target words (*bici* and *casa*) were lowered to 0.33 (LTP words) by adding 12 additional occurrences of their first syllable (i.e., *bi* and *ca*) throughout the corpus. Language B had the same structure but the HTP and LTP words were switched. Although the target words in the corpus were produced in the context of the continuous speech stream, the target word used at test were produced in isolation. In order to ensure that listening preferences were not based on unrelated acoustic differences between the test words, isolated test words were digitally manipulated in Adobe® Audition® to have the same length (500ms) and intensity (65 dB_{SPL}).

Procedure—Experiment 1 consisted of three phases: familiarization (2 min 15 sec), 10-minute delay, and test. Infants were familiarized and tested using the Head Turn Preference Procedure as adapted by Saffran et al. (1996). Infants were seated on a caregiver's lap inside a soundproof booth equipped with a center monitor, two side monitors, and two side audio speakers. The caregiver listened to masking music over headphones to reduce the potential for bias. The experimenter observed the infants' looking behavior in a control room via a closed circuit camera. During familiarization, a video of a flashing light[1] was presented on the monitors contingent upon the infants' looking behavior (as described below in the test phase), while one of the two counterbalanced languages played continuously from the speakers beneath the side monitors.

After the familiarization phase, infants were given a 10-minute break before the testing phase began. During this 10-minute period, infants were allowed to play quietly with toys in the laboratory waiting room, while the caregiver filled out a demographic information questionnaire. Following the 10-minute delay, infants and parents returned to the sound

[1]In the original studies by Pelucchi and colleagues (2009a; 2009b), a flashing bike lights were used as the visual stimulus. However, because our booth was equipped with monitors instead of bike lights, we used a video of the original flashing lights.

booth, for the test phase. All infants heard the same 12 test trials regardless of familiarization condition – each of the four target words (two HTP and two LTP words) were presented three times, randomized by block. Each test trial began with a video of a centrally-presented spinning pinwheel. Since infants were coming back into the booth after a 10-minute delay, we wanted to use a somewhat more engaging visual stimulus to help maintain attention throughout the test phase. Once the infant had fixated the pinwheel, the center monitor was extinguished and the pinwheel appeared on one of the two side monitors. When the infant made a headturn of at least 30 degrees in the direction of side monitor, one of the four isolated target words played continuously, with a 500 ms ISI between target words, until the infant looked away for 2 seconds or until 15 seconds had elapsed. Thus, the infant controlled how long he or she heard the target word. This procedure was repeated until the infant had completed all 12 test trials. Trials with total looking times less that one second were automatically repeated at the end of the test session. The dependent measure was the amount of time the infant oriented toward HTP and LTP words.

Results and Discussion

The difference scores (looking to HTP minus looking to LTP words) from the counterbalanced languages did not vary significantly, $t(22) = .69$, $p = .5$, $d = .29$ (all *t*-tests 2-tailed; effect sizes reported for *t*-tests are Cohen's d), thus, results from the two languages were collapsed in subsequent analyses. A paired samples *t*-test revealed that infants failed to discriminate HTP ($M = 9.82$ s, $SE = .39$) from LTP words ($M = 9.64$ s, $SE = .47$), $t(23) = .58$, $p = .57$, $d = .12$ (see Figure 1), after the 10-minute delay. Failure to discriminate HTP from LTP words following a delay may reflect random or chance performance or instead may reflect similar interest in both HTP and LTP words, possibly due to faster forgetting of HTP relative to LTP words. Either way, the strong co-occurrence information available in the HTP words does not appear to continue to be privileged when infants are tested 10 minutes after familiarization, suggesting that memory for statistical information may decay over even short delays.

The finding that infant memory decays over a short time frame is consistent with recent work by Simon and colleagues (2017) that reported extremely fragile retention of words in a statistical learning paradigm with 6.5-month-old infants. Simon and colleagues (2017) familiarized 6.5-month-old infants with 7 minutes of an artificial language and then tested their ability to discriminated words from part-words followed by a period of sleep or wakefulness. Infants who stayed awake during the delay, showed no evidence of retention. Although infants who napped also failed to show evidence of retention, their performance differed from that of the infants who remained awake, and was correlated with a number of neurophysiological measures. These results suggest that sleep may help promote some retention of fragile representations of statistically defined words.

The word-learning literature also provides evidence of memory decay in older infants (e.g., Horst & Samuelson, 2008; Vlach and Sandhofer, 2012). For example, Horst and Samuelson (2008) found that although 2-year-olds show evidence of having successfully mapped new words to novel referents when tested immediately, they did not retain the word-referent mappings over a 5-minute delay. In another study, Vlach and Sandhofer (2012) found that young children and adults forget newly learned object-label mappings over time. Both the Vlach and Sandhofer (2012) and the Horst and Samuelson (2008) studies further examined whether forgetting may have resulted from insufficient encoding of details of the mapping into memory, by manipulating the encoding conditions. Vlach and Sandhofer (2012) found that providing additional memory support (e.g., shaking the target object, repeating the label multiple times, and asking participants to produce the word for the target object) during the learning phase increases long-term retention of word mappings. Similarly, Horst and Samuelson (2008) showed that children were able to retain words for 5 minutes when their task was augmented with ostensive naming. Together, these studies suggest that providing additional support during learning may have facilitated retention by strengthening the encoding of the relevant words.

Why did the infants in Experiment 1 fail to demonstrate retention of the statistically defined words? In order to efficiently retrieve the statistical properties of words from memory, infants must first robustly encode those patterns. However, brief exposure to a complex natural language may in and of itself not be sufficient to facilitate the robust encoding of words' statistical properties. Without additional support, novel words may not engender such strong representations.

There are a variety of ways in which word representations may be reinforced to support memory. First, infants may benefit from additional exposure to the familiarization corpus. Infants in the current study only heard each of the target words 18 times during approximately 2 minutes of familiarization. Perhaps more experience with the corpus would have allowed infants to more robustly encode the HTP words. It is noteworthy however, that 7 minutes of familiarization with an artificial language was not sufficient for 6.5 month-olds to develop robust representations of statistically defined words (Simon et al., 2017). Second, sleep also appears promote retention of some statistical regularities (e.g., Gomez et al., 2006; Hupbach, et al., 2009; Simon et al., 2017). However, it is not possible to have infants sleep in such a brief 10-minute time frame. A third possibility, and one that we pursue in Experiment 2, is that infants' representations may be strengthened by hearing the target words in isolation. Although the majority of speech infants hear is continuous in nature, a portion of the input infants hear comes in the form of isolated words (Brent & Siskind, 2001; Fernald & Morikawa, 1993; Fernald & Hurtado, 2006). For example, Fernald & Morikawa (1993) and Brent & Siskind (2001) have demonstrated that an average of 9% of utterances produced by caregivers consists of isolated words. These isolated words may play an important role in strengthening infants' lexical representations. Indeed, many of infants' first words are names like Mommy and Daddy, which are also more frequently found in isolation (Ladd, 1997). Further support for this idea comes from recent work by Lew-Williams and colleagues (Lew-Williams et al., 2011). In their study, 8- to 10-month-old infants were familiarized with a somewhat shorter natural Italian corpus that contained either fluent speech only or a combination of fluent speech and isolated words. Infants who heard just the shorter fluent speech stream failed to discriminate HTP from LTP words at test. Infants who heard the same amount of continuous speech but also heard HTP and LTP words presented equally often in isolation at the end of each sentence successfully discriminated the words at test. This work suggests that the combination of isolated words and fluent speech stream appear to support statistical learning. It is important to note that infants heard HTP and LTP words equally often in the corpus and in isolation and so if they had attended only to isolated words, we would expect no difference in looking times for HTP and LTP words. In Experiment 2 we ask whether experience with isolated words also supports memory for statistically defined words.

Experiment 2

Experiment 2 was designed to investigate whether experience with isolated words can selectively reinforce infants' memory for statistically defined words. Twenty-four 8-month-old infants were familiarized with the same Italian corpus as in Experiment 1 and were tested with isolated words both immediately following familiarization (T1) and again following a 10-minute delay (T2). Since language acquisition is fundamentally a dynamic process, we chose to use an infant-controlled procedure to present the isolated words. One of the main advantages of using an infant-controlled paradigm is that it takes into account infants' individual processing and encoding abilities, allowing infants to self-select how long they listen to the target words. This active engagement from the infants might support learning and memory (Perone & Spencer, 2012).

Methods

Participants—Twenty-four healthy, full-term infants (12 males and 12 females) with a mean age of 8.4 months (range = 8 to 8.9) participated in Experiment 2. All other participant characteristics and recruitment methods were identical to Experiment 1. Thirteen additional infants were tested but not included in the analyses for the following reasons: fussiness (n = 9), not paying attention (n = 1), or experimental error (n = 3).

Stimuli—Auditory stimuli were identical to those used in Experiment 1.

Procedures—Procedures were similar to those of Experiment 1, with the following exceptions: 1) infants were tested both immediately after familiarization (T1) and after a 10-minute delay (T2), 2) in order to maintain interest following the 10-minute delay, we used different visual stimuli in the first and second testing sessions. During the familiarization phase and immediate testing phase (T1), a video of flashing lights was presented on the side monitors contingent upon infants looking behavior, similar to Experiment 1. A video of spinning pinwheel was used in place of the video of the flashing light at T2. As in Experiment 1, infants played quietly with toys in laboratory waiting room during the 10-minute delay.

Results and Discussion

We first compared the two counterbalanced familiarization languages for both the immediate testing phase and delayed testing phase. As there was no significant variation in difference scores for the counterbalanced languages at T1, $t(22) = 1.57$, $p = .13$, $d = .64$ or at T2, $t(22) = .62$, p = .54, $d = .26$, the two languages were combined in all subsequent analyses.

Paired samples *t*-test revealed that infants looked significantly longer to HTP words ($M = 9.58$ s, $SE = .46$) than LTP words ($M = 8.6$ s, $SE = .48$) following the 10-minute delay (i.e., T2), $t(23) = 2.81$, $p = .01$, $d = .73$ (see Figure 1). Seventeen out of 24 infants listened longer to the HTP words. In order to explore whether infants in Experiment 2 performed better than those in Experiment 1 (who failed to show discrimination) following a 10-minute delay, we performed a 2 × 2 mixed ANOVA with Word Type (HTP vs LTP) as the within-subjects factor and Experiment (1 vs 2) as the between subjects factor; the interaction between Word Type and Experiment was marginally significant, $F(1, 46) = 3.75$, $p = .059$. Together, results from Experiments 1 and 2 suggest that having experience with isolated HTP and LTP words immediately following familiarization may selectively reinforce memory for HTP words. These findings are consistent with previous research suggesting that experience with isolated words may facilitate statistical learning (Lew-Williams et al., 2011).

Experiment 2 was initially designed to allow infants to control their amount of experience with the isolated words, as we thought that infant memory might benefit more from this type of active engagement (Perone & Spencer, 2012). There are two ways in which this could have supported memory. First, remaining actively engaged immediately following familiarization may have helped support memory for recently segmented words. Second, because infants were able to control how many times they heard the isolated HTP and LTP words at immediate testing (i.e., T1), it is possible that the words that were heard more at T1 and were better remembered following the 10-minute delay (i.e., T2). Thus, if infants preferred listening to the HTP words at T1 – a familiarity preference observed in previous studies using similar materials and procedures (Pelucchi et al, 2009a, 2009b) – the additional experience with the HTP relative to the LTP words may have driven memory for the HTP words following a delay. In this case, the amount of experience with HTP words at T1 should predict the preference for HTP words at T2.

To test this relationship we calculated a differences score (HTP minus LTP) at T1 and T2 and looked at the correlation between these two measures. Importantly, listening preferences at T1 and T2 were not significantly correlated, $r = .11$, $n = 24$, $p = .6$, suggesting that individual differences in experience with the isolated HTP and LTP words at T1 did not impact infants' preferences for HTP words at T2. Thus, although active engagement may still have played a roll in supporting memory, hearing HTP words more at T1 does not appear to predict successful discrimination of HTP versus LTP words at T2.

The lack of correlation between performance at T1 and T2 suggests that infants may not have shown the expected preference for HTP words when tested immediately after familiarization. In fact, a paired *t*-test revealed that infants failed to discriminate HTP words ($M = 7.83$ s, $SE = .44$) from LTP words ($M = 8.29$ s, $SE = .44$), $t(23) = .86$, $p = .4$, $d = .18$ (see Figure 1), when tested immediately after familiarization. This result is surprising given that the familiarization and immediate testing phases from Experiment 2 were essentially an exact replication of Pelucchi et al (2009a, Experiment 3). Failure to replicate previous studies is a pervasive, yet underreported, phenomenon in psychological research (Open Science Collaboration, 2015; see General Discussion for further commentary). Before continuing to pursue to question of the role of isolated words on infants' memory for statistically defined words, we wanted to ensure that we could replicate our original findings when testing infants immediately after familiarization.

Experiment 3

In Experiment 3 we replicate Pelucchi et al (2009a; Experiment 3) and our Experiment 2 T1 (immediate testing) with a new group of 24 eight-month-old infants.

Methods

Participants—Twenty-four healthy, full-term infants (9 males and 15 females) with a mean age of 8.4 months (range = 8 to 8.7) participated in Experiment 3. All other participant characteristics and recruitment methods were identical to Experiments 1 and 2. Seventeen additional infants were tested but not included in the analyses for the following reasons: fussiness (n = 13), not paying attention (n = 2), or experimental error (n = 2).

Stimuli—The auditory stimuli were the same as those used in Experiment 1. In an effort to help infants maintain interest, a video clip of a spinning pinwheel was used instead of a video of a flashing light, during both the familiarization and testing.

Procedures—Experimental procedures were similar to those of Experiment 1 and 2. In Experiment 3, infants were familiarized with the same corpus and were tested immediately following familiarization while watching a video of a spinning pinwheel.

Results and Discussion

As in previous experiments, we first compared the counterbalanced languages. As there were no significant variations between the difference scores for the counterbalanced languages, $t(22) = .22$, $p = .83$, $d = .09$, the two languages were combined in the subsequent analysis. A paired sample *t*-test revealed that infants readily discriminated HTP ($M = 9.74$ s, $SE = .44$) from LTP ($M = 8.87$ s, $SE = .40$) words, $t(23) = 3.68$, $p = .001$, $d = .76$ (see Figure 1). Eighteen out of 24 infants looked longer to HTP words. These results suggest that performance at T1 in Experiment 2 may reflect a Type 2 error (failure to find a true effect). We return to this idea in the General Discussion.

Experiment 4

Results from Experiment 2 suggest that having experience with isolated words immediately following familiarization may reinforce infants' memory for statistically defined words. Infants in Experiment 2 were able to actively select which stimuli they liked to listen to most at both T1 and T2, however, listening preferences following the 10-minute delay were not correlated with listening preferences for HTP versus LTP words immediately after familiarization. These findings suggest that the ability to control their experience with isolated words may not be the primary factor driving improved memory in infants. Rather, it appears that just having experience with isolated words may selectively reinforce memory for HTP words. For a more explicit test of how isolated words support memory, Experiment 4 was designed to replicate and extend the results of Experiment 2 with a more controlled experimental design.

Thus, in Experiment 4, 8-month-olds were given fixed amount of exposure to isolated target words immediately following familiarization with the Italian corpus and were then tested on the same target words following a 10-minute delay. By controlling the amount of experience infants get with the isolated words, we are able to better explore how isolated words may support infant memory for statistically defined words.

Methods

Participants—Twenty-four healthy, full-term infants (12 males and 12 females) with a mean age of 8.3 months (range = 8 to 8.9) participated in Experiment 4. All other participant characteristics and recruitment methods were identical to Experiments 1–3. Seventeen additional infants were tested but not included in the analyses for the following reasons: fussiness (n = 11), not paying attention (n = 3), or experimental error (n = 7).

Stimuli—The auditory stimuli were the same as in the previous experiments. A video clip of a spinning pinwheel was used during both the familiarization and test.

Procedures—Experimental procedures were similar to those of Experiment 2. During familiarization, a video of spinning pinwheel was presented on the monitors contingent upon the infants' looking behavior, while one of the two counterbalanced languages played continuously from the side speakers. Immediately following familiarization, infants were given fixed experience with isolated target words (i.e., infants heard HTP and LTP words equally often in isolation). The two HTP word trials and two LTP word trials were presented three times, randomized by block (i.e., a total of 12 trials). On each trial, a given target word was presented eight times in isolation, for a total of 24 exposures to each target word. This is equivalent to average amount of experience infants received with the isolated words in Experiment 2 at T1.

Following the fixed experience phase, the infant played quietly with the toys in the laboratory waiting room during a 10-minute delay. Infants were tested on their ability to differentiate HTP from LTP words.

Results and Discussion

As in the previous experiments, we first compared the two counterbalanced languages. As there was no significant variation between the difference scores for the counterbalanced languages, $t(22) = 1.13$, $p = .27$, $d = .46$, the two languages were combined in the subsequent analysis. A paired sample *t*-test revealed a significant looking time difference between HTP ($M = 7.47$ s, $SE = .63$) and LTP ($M = 6.52$s, $SE = .53$) words, $t(23) = 2.71$, $p = .012$, $d = .58$ (see Figure 1). Eighteen out of 24 infants looked longer to HTP words. These results suggest that isolated words may play a role in helping infants to successfully encode and remember the statistical properties of words.

General Discussion

The main goal of the current study was to explore the longevity of statistical learning in natural language input by assessing infants' long-term memory for statistically defined words. Infants' ability to retain the statistical properties of words in memory is fundamental to lexical development, especially given that the relevant objects being talked about are not always in the infants' immediate environment. Our findings suggest that infants may initially form weak representations of newly extracted words and that these representations appear to decay over short delays if they are not reinforced. Further, experience with isolated words may function to reinforce initially weak representations, allowing for newly segmented words to be more robustly encoded.

These results may be somewhat surprising in light of previous studies demonstrating relatively strong retention of both words, grammatical regularities, and non-adjacent dependency relationships over delays spanning 5 minutes to 2 weeks (Gómez & Gerken, 1999; Gómez et al., 2006; Hupbach et al., 2009; Jusczyk & Hohne, 1997; Houston & Jusczyk, 2003). There are a number of factors that may account for differences between the current work and previous studies. First, in previous work, infants were provided with much more experience with the relevant stimuli. For example, infants in Jusczyk and Hohne (1997), who recognized the target words following a 2-week delay, heard the target words embedded in sentences for 30 minutes a day for 10 days. Infants in the current student received a short 2 minutes and 15 seconds of exposure to the familiarization language on a single day. Similar, in the work by Gómez and colleagues (Gómez et al., 2006; Hupbach et al., 2009), where infants who napped showed evidence of retaining information about non-adjacent dependency relations 4 and 24 hours after familiarization, infants heard each of the two non-adjacent dependencies 120 times during a 15 minute familiarization phase. In the current study, infants only heard each of the target words 18 times during familiarization. Thus, infants in the current study are likely to have formed a much weaker initial representation of the statistically defined words than fostered in previous work, which may have led to poorer retention.

A second consideration is infants' previous familiarity with the target words. In Houston and Jusczyk's (2003) study, 7.5-month-old infants listened longer to passages containing familiar words (e.g., *cup* and *dog*) that had been heard in isolation 24 hours earlier than to passages containing familiar words (e.g., *feet* and *bike*) not heard on the previous day. Recent work by Bergleson and Swingley (2012, 2015) suggest that infants of this age are likely to already have some sort of representation of the phonological properties of these high frequency words. Thus, exposure to the isolated words during familiarization may have reinforced the already existing phonological representations, allowing Houston and Jusczyk's (2003) infants to recognize those words in continuous speech 24 hours later. Although the Italian words used in the current study were phonotactically legal in English, their phonological realization may not have been familiar to the infants. First, the words themselves were unattested in English and thus infants would not have had preexisting phonological representations of the words. This may also explain the fragile retention of novel statistically defined words documented in 6.5-month-old by Simon and colleagues (2017). Second, a native Italian speaker produced the words and thus their phonological realization would have sounded somewhat unfamiliar to the infants, much like a foreign language or accented speech (see Cristia, Seidl, Vaughn, Schmale, Bradlow, & Floccia, 2012 for a review of processing of accented speech across the lifespan). Thus, in contrast to previous research, infants in the current study may have failed to remember the HTP words following even a short delay interval, because their initial representations were not sufficiently robust to support long-term memory. Future research should investigate whether providing infants with additional experience with the familiarization language or using words that have more overlap with English phonological realizations would better support infants' retention of statistical regularities over time.

Never-the-less, our findings suggest that if infants are exposed the target words in isolation immediately following familiarization, their initially fragile representations can be reinforced sufficiently to support retention over at least short delay intervals. One question that arises from these findings is why would experience with isolated words drive stronger encoding of HTP than LTP words, especially given that all of the isolated words were heard equally often? Had infants just attended to the isolated words in Experiments 2 and 4, we would expect that infants should not differ in their retention of the HTP and LTP words. However, in both experiments, infants showed a significant preference for the HTP following the 10-minute delay, suggesting that they had a more robust memory for the HTP words relative to the LTP words. Successful encoding of the statistical properties of words may require infants to integrate two sources of information: hearing statistically defined words both in the speech stream and in isolation. Integrating information obtained from isolated words with statistical information available in continuous speech may also provide infants with a way to successfully store and eventually retrieve the statistical information from memory.

It is noteworthy that if infants are continually updating co-occurrence statistics, isolated words, both here and in the study by Lew-Williams and colleagues (Lew-Williams et al., 2011), could have functioned to make the task more difficult by reducing the difference in TP between HTP and LTP words. Here, experience with the isolated HTP words did not change the overall TP of those words (i.e., TP remained 1.0). However, by presenting infants with approximately 24 additional tokens of the LTP words immediately after familiarization with the corpus, the between syllable co-occurrence increased in those LTP words (i.e., from a TP of .3 to a TP of .54). This could have functioned to make the LTP words more salient and thus less distinct from the HTP words, however, this is not the pattern that we see here. Of course, it is possible that the transitional probability of HTP (1.0) and LTP (.54) words is sufficiently different to lead to more robust encoding and memory for the HTP words. Further research is needed to determine how infants process and remember words that have graded statistics.

It is also possible that experience with isolated words, on the short term, does not result in an immediate updating of co-occurrence statistics. Instead, once infants have tracked TP information in continuous speech, the representations of sound patterns from the HTP words that are extracted may become available to infants, but only for a very brief amount of time. If those representations are immediately reinforced (here, through experience with isolated words), then they may remain available to infants in long-term memory. Because LTP words may not have been extracted from the corpus initially, infants may fail to have any sort of representation of the LTP words available to be reinforced. Thus, experience with isolated words may selectively benefit infant memory for HTP relative to LTP words, much as experience with isolated words selectively benefited the segmentation of HTP words in the Lew-Williams et al study (2011). The degree to which infants incorporate isolated words into their statistical computations and why isolated words selectively reinforced the HTP words is still unclear. In future work it will be interesting to explore the time frame and conditions over which infants engage in updating co-occurrence patterns.

Although some of the earliest words to appear in infants' lexicons are also found more frequently in isolation, like names (e.g., Mommy, Daddy, infants' own name; Ladd 1997), other common early words include the names of body parts (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012), which do not often occur in isolation (Johnson, Seidl, & Tyler, 2014). Clearly, experience hearing words in isolation is only one of many factors that likely supports infants' ability to segment and remember words. One example of a type of cue that infants may take advantage of in their everyday environment is synchronized touch (e.g., touching the child's knee as the infant hears "there is your knee"). For example, recent work by Seidl and colleagues (Seidl, Tincoff, Baker, & Cristia, 2015) found that synchronous touch cues facilitate the segmentation of words from continuous speech in 4-month-olds. Thus, even when individual words are not often heard in isolation, infant segmentation and memory may be supported not only by the statistical properties of speech, but also by additional cues available in the infant's everyday environment, such as synchronized caregiver touch.

In the process of exploring the role of experience with isolated words on infant memory, in Experiment 2 we failed to replicate the familiarity preference for HTP words seen in Pelucchi et al. (2009a) on immediate test. There are a number of reasons that may account for this failure to replicate. First, it is certainly the case that the present study and Pelucchi et al.'s (2009a, 2009b) original studies were conducted on different populations (i.e., the Midwest vs. the Southeast) and in different laboratories. Further, here we used slightly different visual stimuli (i.e., a video of flashing light instead of an actual flashing light). However, given that infant statistical learning has been demonstrated in many different labs, and using many different types of unrelated visual stimuli (including the same flashing light video and similar pinwheels), it seems unlikely that these differences could account for the discrepancy in the findings. Failure to replicate can also reflect original findings that were erroneous (i.e., Type 1 error). However, we believe our original findings to represent a true effect, as we have demonstrated that infants can track statistical regularities in these naturally produced stimuli across a number of different speakers and corpora (Pelucchi et al., 2009a, 2009b), and again here in Experiment 2 (at T2), Experiment 3, and Experiment 4. Even in studies with high statistical power (i.e., power $\geq .8$), we should expect to fail to replicate ~ 20% of the time – for better or worse, these studies are rarely reported. We too could have abandoned Experiment 2 in the file drawer and just presented Experiments 1 and 4, but fear that doing so may have obfuscated some of the richness in our data. It also would have perpetuated the positive-results bias that is pervasive in psychological literature (Open Science Collaboration, 2015).

In order to explore how accurate our estimation is – that infants can track statistics in natural language input is a true effect – we performed a mini meta-analysis, according to Bergmann and Cristia (2015), on all of our results that used similar natural Italian language materials. As we had all of the relevant data points for our previous studies we were able to calculate Cohen's d and t-values for use in the meta-analysis. The mini meta-analysis was performed in R (R Core Team, 2013), using a MetaLab script as provided by Bergmann and colleagues

(Bergmann, Tsuji, Piccinini, Lewis, Braginsky, Cristia, & Frank, in prep). As seen in the resulting Forest Plot (see Figure 2), results from the immediate testing phase (T1) of Experiment 2 clearly fall outside of the confidence interval for comparable studies. Further, results of Experiment 2 at T1 are not weighted heavily in the meta-analysis (as reflected in the small size of the point/box for this study). This suggests that there is considerable noise in the measurement. Together, these primary outcomes measures of the meta-analysis suggest that results from this particular experiment likely reflect expected noise (or ~20% of studies with a Type 2 error), and thus performance in Experiment 2 at T1 might be largely due to expected variance in sampling.

While it is theoretically possible that infants' failure to discriminate HTP from LTP words following a 10-minute delay in Experiment 1 also reflect a Type 2 error, the mini meta-analysis suggests otherwise. Specifically, the results of Experiment 1 pattern with a subset of other studies where infants were not predicted to be successful at discriminating HTP from LTP words. In Lew-Williams et al. (2011) Experiment 1a (fluent speech only condition) infants were provided with so little exposure to the Italian corpus (i.e., HTP and LTP words were only heard 12 times each and were embedded in the fluent speech stream), that infants were not expected to successfully segment the speech stream. Indeed, infants did not discriminate HTP from LTP words at test. In Lew-Williams et al. (2011) Experiment 2 infants were presented with the same amount of fluent speech but also heard unrelated isolated words presented through the corpus. Again, as expected, infants failed to discriminate HTP from LTP words at test, suggesting that unrelated isolated words do not support statistical learning in the same way that related isolated words do. All three of these studies (i.e., Experiment 1, Lew-Williams et al. Experiment 1a and Lew-Williams et al. Experiment 2), are weighted fairly similarly in the meta-analysis and have a relatively low amount of noise in their measurement. Thus, a more parsimonious explanation for the results of Experiment 1 is that infants' memory for the statistical defined words truly decayed over the 10-minute retention interval.

Meta-analytical approaches provide a powerful tool for estimating the effect size and its variance across a number of studies investigating the same phenomenon. As outlined by Bergmann and Cristia (2015), the meta-analytic approach can go well beyond this initial scientific goal. It can also be used to assess the impact of factors of interest including methodological factors and participant characteristic on experimental outcomes. Further, the meta-analytic approach can provide guidance for future research in helping to inform experimental decisions such as appropriate sample size. Bergmann and Cristia (2015) provide a succinct and helpful review of the benefits of and best practices in using a meta-analytic approach for the interested reader (see also Bergmann et al, in prep; Borenstein, Hedges, Higgins, & Rothstein, 2009; Lewis et al., 2017).

There are a number of other avenues of future research that may provide evidence for the role of statistical learning during natural language acquisition. First, it will be interesting to investigate whether sleep promotes infants' memory for statistically defined words by allowing their brain to organize and consolidate memory traces. Sleep and wakeful rest have been implicated in memory consolidation (e.g., Stickgold, 2005; Stickgold & Walker, 2005) and work by Gómez and colleagues (Gómez et al., 2006; Hupbach et al., 2009) suggests that, for 15-month-olds, sleep in the form of naps promotes the consolidation and abstraction of newly learned simple non-adjacent dependency relations (see also Simon et al., 2017). A second line of work, and one that we are currently pursuing, is exploring whether the statistically defined words that infants and young children segment from a fluent speech stream can function as candidate object labels in a novel word learning task implemented following a delay. This work will shed light on the robustness of infants' representations of newly segmented words and the relevance of statistical learning to language acquisition.

Together, our findings suggest that the experience of hearing words in isolation immediately after familiarization selectively reinforces infants' long-term memory for high transitional probability words. These findings add significantly to the existing knowledge on infant statistical learning and provide initial support for the longevity of statistical learning in the first year. Memory for statistically defined words could bolster language acquisition by allowing those sound sequences to have a privileged status in new learning environments.

Acknowledgments

This research was funded in part by a grant from NICHD to JFH (R01HD083312) and a fellowship from Turkish Ministry of National Education to FK. We would like to thank the members of the Infant Language and Perceptual Learning Lab, the participating families, Christina Bergmann for her help in generating the forest plot and conducting and interpreting the mini meta-analysis, and three anonymous reviewers and the editor for helpful comments on a previous version of this manuscript.

References

Akhtar N. The robustness of learning through overhearing. Developmental Science. 2005; 8(2):199–209. DOI: 10.1111/j.1467-7687.2005.00406.x [PubMed: 15720377]

Akhtar N, Jipson J, Callanan M. Learning words through overhearing. Child Development. 2001; 72(2):416–430. DOI: 10.1111/1467-8624.00287 [PubMed: 11333075]

Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. Psychological Science. 1998; 9(4):321–324. DOI: 10.1111/1467-9280.00063

Bauer PJ. Constructing a past in infancy: A neuro-developmental account. Trends in Cognitive Sciences. 2006; 10(4):175–181. DOI: 10.1016/j.tics.2006.02.009 [PubMed: 16537115]

Bergelson E, Swingley D. At 6 to 9 months, human infants know the meanings of many common nouns. Proceedings of the National Academy of Sciences of the USA. 2012; 109(9):3253–3258. DOI: 10.1073/pnas.1113380109 [PubMed: 22331874]

Bergelson E, Swingley D. Early word comprehension in infants: Replication and extension. Language Learning and Development. 2015; 11(4):369–380. DOI: 10.1080/15475441.2014.979387 [PubMed: 26664329]

Bergmann C, Cristia A. Development of infants' segmentation of words from native speech: a meta-analytic approach. Developmental Science. 2015; 19(6):901–917. DOI: 10.1111/desc.12341 [PubMed: 26353859]

Bergmann C, Tsuji S, Piccinini P, Lewis M, Braginsky M, Cristia A, Frank M. Planning for reproducibility based on previous infant research. 2017 Manuscript in preparation.

Brent MR, Siskind JM. The role of exposure to isolated words in early vocabulary development. Cognition. 2001; 81(2):B33–B44. DOI: 10.1016/S0010-0277(01)00122-6 [PubMed: 11376642]

Bosch L, Figueras M, Teixidó M, Ramon-Casas M. Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages. Frontiers in Psychology. 2013; 4:106.doi: 10.3389/fpsyg.2013.00106 [PubMed: 23467921]

Cole, R., Jakimik, J. A model of speech perception. In: Cole, R., editor. Perception and Production of Fluent Speech. Hillsdale, NJ: Erlbaum; 1980. p. 133-163.

Collie R, Hayne H. Deferred imitation by 6-and 9-month-old infants: More evidence for declarative memory. Developmental Psychobiology. 1999; 35(2):83–90. DOI: 10.1002/(SICI)1098-2302(199909)35:2<83::AID-DEV1>3.0.CO;2-S [PubMed: 10461122]

Cowan N, Elliott EM, Saults JS, Morey CC, Mattox S, Hismjatullina A, Conway AR. On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. Cognitive Psychology. 2005; 51(1):42–100. DOI: 10.1016/j.cogpsych.2004.12.001 [PubMed: 16039935]

Estes KG, Evans JL, Alibali MW, Saffran JR. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. Psychological Science. 2007; 18(3):254–260. DOI: 10.1111/j.1467-9280.2007.01885.x [PubMed: 17444923]

Fernald A, Hurtado N. Names in frames: Infants interpret words in sentence frames faster than words in isolation. Developmental Science. 2006; 9(3):F33–F40. DOI: 10.1111/j.1467-7687.2006.00482.x [PubMed: 16669790]

Fernald A, Morikawa H. Common themes and cultural variations in Japanese and American mothers' speech to infants. Child Development. 1993; 64(3):637–656. DOI: 10.1111/j.1467-8624.1993.tb02933.x [PubMed: 8339686]

Friederici AD, Wessels JM. Phonotactic knowledge of word boundaries and its use in infant speech perception. Attention, Perception, & Psychophysics. 1993; 54(3):287–295. DOI: 10.3758/BF03205263

Goyet L, Nishibayashi LL, Nazzi T. Early syllabic segmentation of fluent speech by infants acquiring French. PloS ONE. 2013; 8(11):e79646.doi: 10.1371/journal.pone.0079646 [PubMed: 24244536]

Gómez RL. Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. Philosophical Transactions of Royal Society B. 2017; 372(1711):20160054.doi: 10.1098/rstb.2016.0054

Gómez RL, Bootzin RR, Nadel L. Naps promote abstraction in language-learning infants. Psychological Science. 2006; 17(8):670–674. DOI: 10.1111/j.1467-9280.2006.01764.x [PubMed: 16913948]

Hay JF, Pelucchi B, Estes KG, Saffran JR. Linking sounds to meanings: infant statistical learning in a natural language. Cognitive Psychology. 2011; 63(2):93–106. DOI: 10.1016/j.cogpsych.2011.06.002 [PubMed: 21762650]

Hay JF, Saffran JR. Rhythmic grouping biases constrain infant statistical learning. Infancy. 2012; 17(6):610–641. DOI: 10.1111/j.1532-7078.2011.00110.x [PubMed: 23730217]

Herbert J, Hayne H. Memory retrieval by 18–30-month-olds: Age-related changes in representational flexibility. Developmental Psychology. 2000; 36(4):473.doi: 10.1037/0012-1649.36.4.473 [PubMed: 10902699]

Houston DM, Jusczyk PW. Infants' long-term memory for the sound patterns of words and voices. Journal of Experimental Psychology: Human Perception and Performance. 2003; 29(6):1143.doi: 10.1037/0096-1523.29.6.1143 [PubMed: 14640835]

Houston DM, Jusczyk PW, Kuijpers C, Coolen R, Cutler A. Cross-language word segmentation by 9-month-olds. Psychonomic Bulletin & Review. 2000; 7(3):504–509. DOI: 10.3758/BF03214363 [PubMed: 11082857]

Horst JS, Samuelson LK. Fast mapping but poor retention by 24-month-old infants. Infancy. 2008; 13(2):128–157. DOI: 10.1080/15250000701795598

Höhle B, Weissenborn J. German-learning infants’ ability to detect unstressed closed-class elements in continuous speech. Developmental Science. 2003; 6(2):122–127. DOI: 10.1111/1467-7687.00261

Hunter M, Ames E. A multifactor model of infant preferences for novel and familiar stimuli. Advances in Infancy Research. 1988; 5:69–95.

Hupbach A, Gómez RL, Bootzin RR, Nadel L. Nap-dependent learning in infants. Developmental Science. 2009; 12(6):1007–1012. DOI: 10.1111/j.1467-7687.2009.00837.x [PubMed: 19840054]

Johnson EK, Seidl A, Tyler MD. The edge factor in early word segmentation: utterance-level prosody enables word form extraction by 6-month-olds. PloS ONE. 2014; 9(1):e83546.doi: 10.1371/journal.pone.0083546 [PubMed: 24421892]

Johnson EK, Tyler MD. Testing the limits of statistical learning for word segmentation. Developmental Science. 2010; 13(2):339–345. DOI: 10.1111/j.1467-7687.2009.00886.x [PubMed: 20136930]

Jusczyk PW, Aslin RN. Infants’ detection of the sound patterns of words in fluent speech. Cognitive Psychology. 1995; 29(1):1–23. DOI: 10.1006/cogp.1995.1010 [PubMed: 7641524]

Jusczyk PW, Cutler A, Redanz NJ. Infants' preference for the predominant stress patterns of English words. Child Development. 1993; 64(3):675–687. DOI: 10.1111/j.1467-8624.1993.tb02935.x [PubMed: 8339688]

Jusczyk PW, Hohne EA. Infants' memory for spoken words. Science. 1997; 277(5334):1984–1986. DOI: 10.1126/science.277.5334.1984 [PubMed: 9302291]

Jusczyk PW, Hohne EA, Bauman A. Infants’ sensitivity to allophonic cues for word segmentation. Attention, Perception, & Psychophysics. 1999; 61(8):1465–1476. DOI: 10.3758/BF03213111 [PubMed: 10598463]

Jusczyk PW, Houston DM, Newsome M. The beginnings of word segmentation in English-learning infants. Cognitive Psychology. 1999; 39(3):159–207. DOI: 10.1006/cogp.1999.0716 [PubMed: 10631011]

Ladd, R. Intonational phonology. London: Cambridge University Press; 1997.

Lewis M, Braginsky M, Tsuji S, Bergmann C, Piccinini PE, Cristia A, Frank MC. A quantitative synthesis of early language acquisition using meta-analysis. 2017 Retrieved from osf. io/htsjm.

Lew-Williams C, Pelucchi B, Saffran JR. Isolated words enhance statistical language learning in infancy. Developmental Science. 2011; 14(6):1323–1329. DOI: 10.1111/j.1467-7687.2011.01079.x [PubMed: 22010892]

Mattys SL, Jusczyk PW. Phonotactic cues for segmentation of fluent speech by infants. Cognition. 2001; 78(2):91–121. DOI: 10.1016/S0010-0277(00)00109-8 [PubMed: 11074247]

Meltzoff AN. Immediate and deferred imitation in fourteen- and twenty-month-old infants. Child Development. 1985; 56(1):62–72. DOI: 10.2307/1130174

Meltzoff AN. Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. Developmental Psychology. 1988; 24(4):470–476. DOI: 10.1037/0012-1649.24.4.470 [PubMed: 25147404]

Mersad K, Nazzi T. When Mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. Language Learning and Development. 2012; 8(3):303–315. DOI: 10.1080/15475441.2011.609106

Nakashiba T, Young JZ, McHugh TJ, Buhl DL, Tonegawa S. Transgenic inhibition of synaptic transmission reveals role of CA3 output in hippocampal learning. Science. 2008; 319(5867):1260–1264. DOI: 10.1126/science.1151120 [PubMed: 18218862]

Nyberg, L., Cabeza, R. Brain imaging of memory. In: Tulving, E., Craik, FIM., editors. The Oxford handbook of memory. New York: Oxford University Press; 2000. p. 501-519.

Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015; 349(6251):aac4716.doi: 10.1126/science.aac4716 [PubMed: 26315443]

Pelucchi B, Hay JF, Saffran JR. Statistical Learning in a Natural Language by 8-Month-Old Infants. Child Development. 2009; 80(3):674–685. DOI: 10.1111/j.1467-8624.2009.01290.x [PubMed: 19489896]

Pelucchi B, Hay JF, Saffran JR. Learning in reverse: Eight-month-old infants track backward transitional probabilities. Cognition. 2009; 113(2):244–247. DOI: 10.1016/j.cognition.2009.07.011 [PubMed: 19717144]

Perone S, Spencer JP. Autonomy in action: Linking the act of looking to memory formation in infancy via dynamic neural fields. Cognitive Science. 2013; 37(1):1–60. DOI: 10.1111/cogs.12010 [PubMed: 23136815]

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2013. URL http://www.R-project.org/

Rosenblad A. Introduction to Meta-Analysis by Michael Borenstein, Larry V. Hedges, Julian P.T. Higgins, Hannah R. Rothstein. International Statistical Review. 2009; 77:478–479. DOI: 10.1111/j.1751-5823.2009.00095_15.x

Ross-Sheehy S, Newman RS. Infant auditory short-term memory for non-linguistic sounds. Journal of Experimental Child Psychology. 2015; 132:51–64. DOI: 10.1016/j.jecp.2014.12.001 [PubMed: 25590900]

Ross-sheehy S, Oakes LM, Luck SJ. The development of visual short-term memory capacity in infants. Child Development. 2003; 74(6):1807–1822. DOI: 10.1046/j.1467-8624.2003.00639.x [PubMed: 14669897]

Rovee-Collier C. The development of infant memory. Current Directions in Psychological Science. 1999; 8(3):80–85. DOI: 10.1111/1467-8721.00019

Rovee CK, Fagen JW. Extended conditioning and 24-hour retention in infants. Journal of Experimental Child Psychology. 1976; 21(1):1–11. DOI: 10.1016/0022-0965(76)90052-7 [PubMed: 1270997]

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science. 1996; 274(5294):1926–1928. DOI: 10.1126/science.274.5294.1926 [PubMed: 8943209]

Seidl A, Tincoff R, Baker C, Cristia A. Why the body comes first: effects of experimenter touch on infants' word finding. Developmental Science. 2015; 18(1):155–164. DOI: 10.1111/desc.12182 [PubMed: 24734895]

Simon KN, Werchan D, Goldstein MR, Sweeney L, Bootzin RR, Nadel L, Gómez RL. Sleep confers a benefit for retention of statistical language learning in 6.5 month old infants. Brain and Language. 2017; 167:3–12. DOI: 10.1016/j.bandl.2016.05.002 [PubMed: 27291337]

Stickgold R. Sleep-dependent memory consolidation. Nature. 2005; 437(7063):1272–1278. DOI: 10.1038/nature04286 [PubMed: 16251952]

Stickgold R, Walker M. Memory consolidation and reconsolidation: What is the role of sleep. Trends in Neurosciences. 2005; 28(8):408–415. DOI: 10.1016/j.tins.2005.06.004 [PubMed: 15979164]

Squire LR. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychological Review. 1992; 99:195–231. DOI: 10.1037/0033-295X.99.2.195 [PubMed: 1594723]

Thiessen ED, Saffran JR. When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Developmental Psychology. 2003; 39(4):706.doi: 10.1037/0012-1649.39.3.484 [PubMed: 12859124]

Tincoff R, Jusczyk PW. Some beginnings of word comprehension in 6-month-olds. Psychological Science. 1999; 10(2):172–175. DOI: 10.1111/1467-9280.00127

Tincoff R, Jusczyk PW. Six-Month-Olds comprehend words that refer to parts of the body. Infancy. 2012; 17(4):432–444. DOI: 10.1111/j.1532-7078.2011.00084.x

Vlach HA, Sandhofer CM. Fast mapping across time: Memory processes support children's retention of learned words. Frontiers in Psychology. 2012; 3(4):46.doi: 10.3389/fpsyg.2012.00046 [PubMed: 22375132]

Wojcik EH. Remembering new words: integrating early memory development into word learning. Frontiers in Psychology. 2013; 4:151.doi: 10.3389/fpsyg.2013.00151 [PubMed: 23554599]

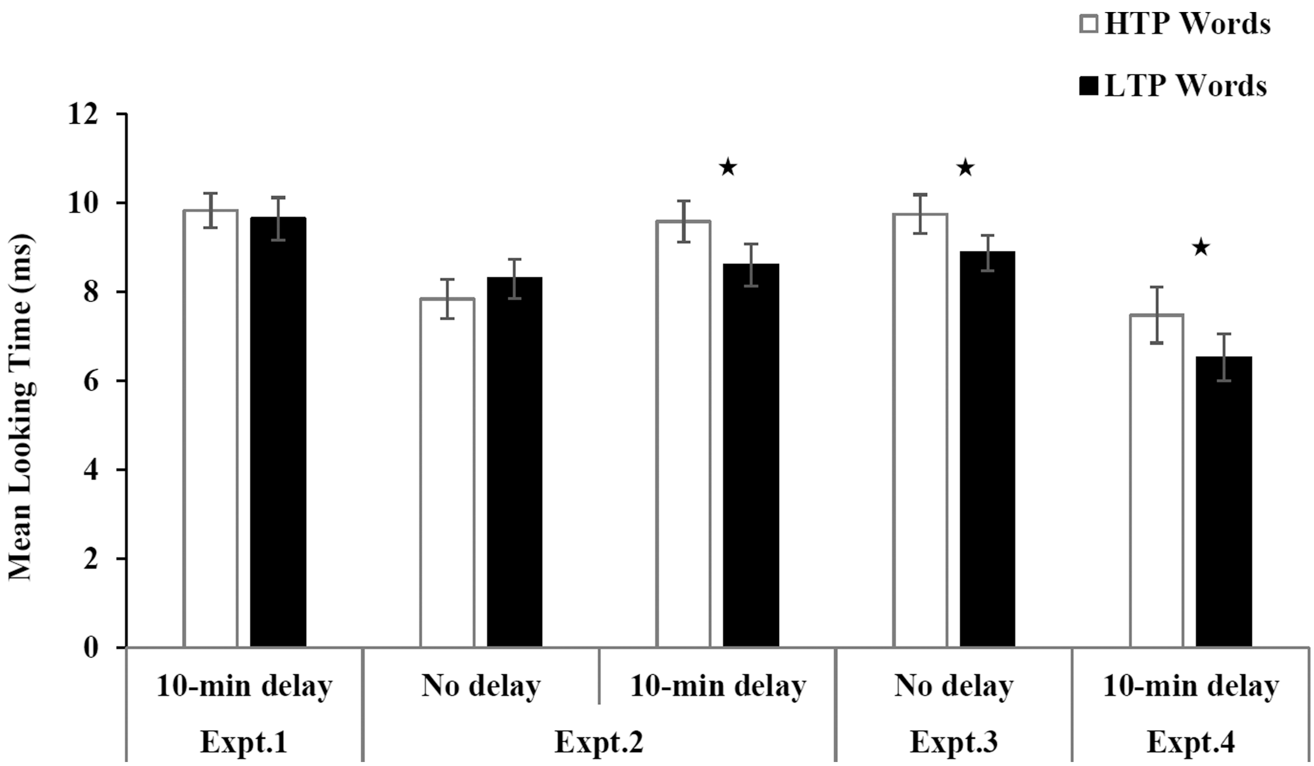

Figure 1.

Mean looking times (ms) to HTP and LTP words. Error bars represent standard error of the mean.

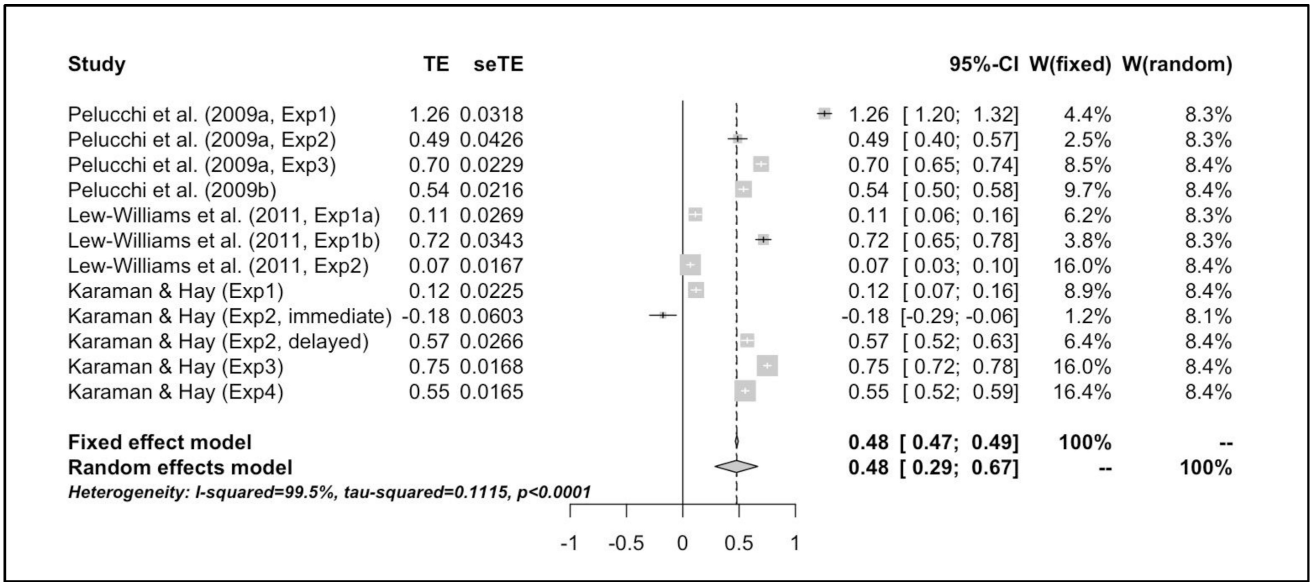

Figure 2.

Forest plot displaying results from a mini meta-analysis of the word segmentation studies that used our natural Italian corpora. The fixed-effects model assumes that we have sampled from a single true effect. The random-effects model assumes that we have sampled from a distribution of effects. Although we present both the fixed- and random-effects model here, given our small sample size and our belief that we are sampling from a distribution of effects, best practices suggest focusing on the random-effects model (Bergmann & Cristia, 2015; Rosenblad, 2009). Each study is represented by a line in the plot. TE: Treatment Effect (Effect Size); seTE: Estimated Treatment Effect; CI: Confidence Interval; Diamond: Overall Effect Estimate; Width of diamond: CI for overall effect estimate; Point/box size: Inverse variance weight based on effect size SE; Width of line: CI for effect estimate for each study – lines are plotted in black if the CI is larger than that point/box size and in white if the CI falls within the point/box.