Infant Statistical-Learning Ability is Related to Real-Time Language Processing

Jill Lany, Amber Shoaib, Abbie Thompson

Department of Psychology

University of Notre Dame

and

Katharine Graf Estes

Department of Psychology

University of California-Davis

Address correspondence to:

Jill Lany

Department of Psychology

Haggar Hall

Notre Dame, IN 46556

Email: jlany@nd.edu

Acknowledgements: This work was supported by funds from NSF BCS-1352443and by Notre Dame's Institute for Scholarship in the Liberal Arts to J.L.

Abstract

Infants are adept at learning statistical regularities in artificial language materials, suggesting that the ability to learn statistical structure may support language development. Indeed, infants who perform better on statistical learning tasks tend to be more advanced in parental reports of infants' language skills. Work with adults suggests that one way statistical learning ability affects language proficiency is by facilitating real-time language processing. Here we tested whether 15-month-olds' ability to learn sequential statistical structure in artificial language materials is related to their ability to encode and interpret native-language speech. Specifically, we tested their ability to learn sequential structure among syllables (Experiment 1) and words (Experiment 2), as well as their ability to encode familiar English words in sentences. The results suggest that infants' ability to learn sequential structure among syllables is related to their lexical-processing efficiency, providing continuity with findings from children and adults, though effects were modest.

Spoken language is a part of virtually all hearing children's daily experience, and they clearly learn what they hear. The words that are most prevalent in children's language input are learned relatively early (Huttenlocher et al., 1991), and most of children's early multi-word utterances are imitations of adult productions (Lieven et al., 2003). However, beyond learning specific words and phrases from the speech they hear, infants may also be learning about the statistical patterns of their language. There are many instances of sequential statistical structure in spoken language, such as the fact that syllables that reliably co-occur are likely to belong to the same word, while syllables that rarely co-occur are more likely to span word boundaries (e.g., Swingley, 2005). Beyond syllable-level statistics, words are ordered in highly predictable ways in many languages. For example, in English determiners tend to co-occur with nouns, while pronouns and auxiliaries tend to co-occur with verbs (e.g., Mintz et al., 2002).

Even very young infants can learn novel sequential statistics in artificial languages with just a few minutes of exposure (see Lany & Saffran, 2013, for a review). Much work has focused on infants' ability to learn a form of sequential structure referred to as a transitional probability, or TP. The TP between two elements, X and Y, is computed by dividing the frequency of XY by the frequency of X, yielding the probability that if X occurs, Y will also occur. Transitional probabilities between syllables tend to be higher within words than between them, and thus provide information that is relevant to segmenting word forms within fluent speech. Infants can track TPs between syllables in synthesized and naturally produced speech, suggesting that it may be one mechanism by which word forms are segmented (e.g., Aslin et al., 1998; Saffran et al., 1996; Pelucci et al, 2009). Furthermore, infants map sequences with high internal TPs to referents, but do not do so for sequences that are equally frequent but have low internal TPs (Graf Estes et al., 2007; Graf Estes, 2012; Hay et al., 2011). The fact that infants selectively treat high TP sequences as labels suggests that sequential learning can yield word-like representations, and that such learning is potentially relevant to lexical development.

Infants can also track sequential dependencies at higher levels of language structure, such as the likelihood that words or word categories will co-occur (e.g., determiner-noun and auxiliary-verb co-occurrence relationships). Many studies have used an *aX bY* artificial language, in which *a*s are followed by *X*s and *b*s by *Y*s but not vice versa, to study this sequential structure. Infants learn these co-occurrence relationships best when *X*s and *Y*s can be distinguished by salient perceptual cues, such as differences in their phonological characteristics (Gerken et al., 2005; Gómez & Lakusta, 2004; Lany & Gómez, 2008). By 12 months of age infants can learn the *aX bY* structure, even when it is probabilistic (Gómez & Lakusta, 2004). Moreover, by 22 months of age infants can use the correlated phonological and distributional cues in the *aX bY* language to help them form robust word-referent associations, and to guide their interpretation of novel words when reference is ambiguous (Lany & Saffran, 2010). Specifically, infants were given auditory experience with an *aX bY* language as described above (the

Experimental Group), or with a version of the language in which distributional and phonological cues did not reliably cue category membership (the Control group). Both groups were then trained on pairings between the languages' word-categories and pictures of animals and vehicles (e.g., *aX*-phrases labeled animals, and *bY*-phrases labeled vehicles), and tested on trained and novel picture-word pairings. Only Experimental infants learned the associations between words and pictures they were trained on and generalized the pattern to novel pairings, suggesting that statistical cues were extremely important to word learning. These findings parallel those from studies of native language development, which show that children use distributional cues to help them establish whether words refer to objects or actions by 24-onths of age (Waxman et al., 2009).

These studies suggest that infants are sensitive to sequential statistical information that has *potential* to be highly informative for acquiring their native language. A handful of studies have directly examined whether infants' ability to detect to novel sequential statistics presented in the auditory modality is related to native language development, typically their scores on the MacArthur-Bates Communicative Development Inventory (MCDI), which is a normed parent-report measure of vocabulary size. Some researchers have approached this question by testing whether infants' ability to learn phonotactic patterns that are not allowed in their native language is related to their native language proficiency. For example, 19 month-old infants with larger vocabularies were less likely to learn phonotactic patterns that are illegal in English (i.e., words beginning with the sound sequences *tl*, *ps, fw, and shn)* than infants who knew fewer English words (Graf Estes, Gluck, & Grimm, 2016). Infants with larger vocabularies are also less likely to learn words with illegal phonotactic patterns as object labels (Graf Estes, Edwards, & Saffran, 2011). These findings suggest that infants who know more words in their native language tend to be more entrenched in their native language phonotactics, resistant to learning patterns that do not conform to them.

Another line of work has investigated the ability to learn statistical regularities in nonlinguistic materials, thus avoiding the problem of native language interference. For example, Shafto et al. (2012) tested whether 8.5 month-olds' ability to learn a visuo-spatial sequence was related to concurrent and later parent-report measures of language development. Infants were presented with triplets of shapes, with each triplet occurring in a consistently ordered spatial configuration. Infants' visual sequence learning was related to both gesture use and receptive vocabulary size, though the effects were modest and somewhat inconsistent across the ages tested (see also Ellis et al., 2013).

In a related study Kidd and Arciuli (2016) tested whether 6-8 years olds' ability to learn nonlinguistic visual sequences was related to their performance on a sentence comprehension task. The statistical learning task tapped children's ability to learn high TP triplets within a series of cartoon-like aliens. The sentence comprehension task involved picking the picture described by a sentence (i.e., “Which chicken is being kissed by the mouse?”). Children who exhibited better visual sequence learning also demonstrated better comprehension of the most challenging sentences, even when controlling for measures of their working memory capacity and nonverbal IQ. These findings provide compelling evidence that statistical learning ability is related to language proficiency in childhood. The work is also notable in its use of a behavioral test to measure specific aspects of native-language proficiency, rather than more general language outcome measures.

Other studies have approached the task of investigating relations between statistical learning and language development by using artificial languages that contain statistical regularities that are novel, but that do not directly violate native language structures. For example, Lany tested whether infants who are better able to capitalize on statistical regularities in an *aX bY* language, which is modeled after determiner-noun and auxiliary-verb co-occurrence relations in English, are also more advanced in their native language development (Lany, 2014; Lany & Safffran, 2011). By 22 months of age infants use the statistical structure in the language to help them learn word-referent mappings (Lany & Saffran, 2010). Importantly, infants who are better able to use probabilistic sequential regularities to learn word-referent mappings in an artificial language are more likely to have begun combining words regularly in their own speech (Lany, 2014). In addition, even though infants benefit from both distributional and phonological cues in these tasks, infants with larger vocabularies are more likely to rely on word-order cues to generalize (i.e., using the determiner-like word to help establish reference), while infants with smaller vocabularies are more likely to use the phonological properties of the words to do so (Lany & Saffran, 2011).

One advantage to using artificial languages that are constructed to parallel natural language is that it allows researchers to test connections between native language development and learning statistical regularities that clearly tap processes relevant to language learning. For example, by studying how infants use statistical regularities to learn words, the studies described above may shed light on developmental differences in the accessibility of different kinds of cues in word learning tasks (Lany, 2014; and Lany & Saffran, 2011). We note that there is convergence across tasks, as the relations observed in studies using artificial languages modeled after natural languages are consistent with those from work examining relations between tasks assessing statistical learning in the visual modality across infancy and childhood (Shafto et al., 2012; Kidd 2016) as well as with the relations between infants' knowledge of native language statistics and vocabulary development (Graf Estes et al., 2016).

Altogether, the studies reviewed above provide empirical support for the hypothesis that infants' ability to learn sequential structure plays a role in language development. Specifically, they suggest that performance on tasks assessing sequential statistical learning are related to parent-report measures of language development, such as vocabulary size and grammatical development. They are also consistent with evidence that sensitivity to sequential statistical structure is related to real-time language-language processing in adults. For example, Farmer et al. (2006) tested how quickly adults recognize nouns as a function of their phonological properties. The target words occurred in sentence contexts that provided sequential cues that unambiguously suggested the next word should be a noun (e.g., "The curious boy saved the ____"). Participants were faster to recognize target words when they had phonological characteristics that were more typical of nouns (e.g., "marble") than when their phonology was more typical of verbs (e.g., "insect"). Likewise, Conway et al. (2010) found that adults who were better able to learn novel sequential patterns in artificial grammars were more likely to capitalize on sequential structure to recognize words in native language speech.

Conway et al. (2010) interpret these findings as evidence that individuals with better statistical learning abilities are better able to learn sequential structure in their native language, and in turn are able to capitalize on such structure to facilitate encoding and interpreting speech. However, we do not know whether statistical learning skill is related to speech processing during language acquisition. Infants' lexical-processing efficiency, or LPE, can be tested by presenting a familiar "target" word (e.g., "Find the doggie!") in conjunction with a picture array containing the target referent (i.e., a dog) and a distractor (e.g., a baby). Relative to adults, infants and children have poor LPE, recognizing words more slowly and less accurately, but LPE improves substantially across the second year of life (Fernald et al., 1998 2001). It is possible that infants' ability to track sequential statistics facilitates gains in LPE.

Thus, in the current work we tested the hypothesis that infants' sequential statistical learning ability is related to LPE across two experiments using well-studied artificial languages that contain sequential structure. In both experiments we assessed LPE using the task described above, which has been used extensively to study individual differences in real-time processing in infants of the ages we tested here (e.g., Fernald et al. 2006; 2008). In Experiment 1 we tested whether infants' performance on a word segmentation task is related to their performance on the LPE task. In Experiment 2 we tested whether infants' ability to track word-order patterns is related to their LPE.

Experiment 1

We asked whether infants who are better able to learn sequential structure in speech are better able to recognize word forms quickly in real time. We tested statistical learning using an artificial language that was composed of a continuous stream of syllables. Within the stream there were syllable sequences with both High and Low TPs (Graf Estes et al., 2007). We tested whether infants were able to track the TPs by asking whether they discriminate between the HTP and LTP syllable sequences using the Head-Turn Preference Procedure (HTPP). In the HTTP, infants' ability to discriminate between two types of stimuli is assessed through preferential listening. Learning can be expressed both as a novelty preference and a familiarity preference, depending on many factors, such as the social relevance of the stimuli, and infants' age and encoding ability (Hunter & Ames, 1998; Houston-Price & Nakia, 2004).

The HTPP is the task most commonly used to study infant statistical learning in the auditory modality. Thus, there is a great deal of value in using this task to measure statistical learning in the current work. The HTPP has predominantly been used to test language learning and development by examining group-level performance, however it has recently been used to test whether there is a connection between individual infants' speech segmentation ability and standardized assessments of their language development (Newman et al, 2006; Newman et al, 2016; Singh et al., 2012; see also Graf Estes et al., 2016). Thus, the HTPP may also reveal meaningful individual differences in sensitivity to TPs relevant for word segmentation in the current experiment.

Furthermore, by examining whether there are individual differences in infants' preference patterns (i.e., novelty versus familiarity, and magnitude of preference), this work may shed light on failures to replicate prior findings. In developmental research with the HTPP, researchers typically focus on tight age ranges in an effort to reduce individual variability, but age is not necessarily the best predictor of language development. Thus, including characteristics of the sample that pertain to language development (e.g., LPE) may help researchers determine whether differences in results are related to differences in language skills across samples, or to create groups based on age and language competence rather than on age alone.

To measure LPE, we used a task designed by Fernald and colleagues that assesses infants' ability to use spoken language to find matching visual objects (see Fernald et al., 2008, for an overview of the method). On a given trial, infants might view a picture of a dog and a picture of a baby, and hear "Find the baby!". The words selected for testing are highly likely to be known by infants and toddlers at this age, and thus our participants were likely to be able to find the correct picture. Of primary interest is the speed with which they did so. We chose this task because it has been used extensively (e.g., Fernald et al., 1998; Fernald et al., 2006), and because infants with larger vocabularies at 15 months tend to process speech more rapidly by the time they are about 2 years of age (Fernald et al., 2006).

We tested 15-month-old infants because this is the youngest age at which Fernald recommends using the LPE task (Fernald et al., 2008). The artificial language we chose had not previously been used with 15-month-olds. Thus, we did not have an a priori prediction for the direction of preference that would be associated with better statistical learning. Rather, we predicted that the magnitude of infants' preferences would be associated with their performance on the LPE task.

Methods

Participants. Participants were 38 English-learning infants between 15.0 and 16.0 months of age (21 were female; see Table 1 for additional descriptive statistics). Infants were free of hearing, vision, or language-development problems based on parental report. Infants were not eligible if they were born before 36 weeks gestation, weighed less than 5lbs 5oz at birth. An additional 16 infants were tested but their data were not included due to failure to contribute sufficient data on the LPE (n = 5) or statistical learning tasks (n = 5), due to fussiness or inattention, equipment failure or experimenter error (n = 4), falling asleep (n = 1), or because their scores on any of the measures were greater than 3 standard deviations from the mean (n = 1).[1] Parental consent was obtained for all participants. Participants were given a children's book or 15 dollars for participating in the study.

Materials and Procedure. Infants were first trained and tested on the statistical learning task, followed by the LPE task. A parent or caregiver accompanied the infant for the duration of the experiment, and they were instructed not to talk during that time. Both tasks took place in an 8' X 8' sound-attenuated room. A 60" LCD screen was mounted on the central wall, and a 20" monitor with a speaker behind it was mounted on each side wall, all at the height of 31". A chair for the parent was located approximately 3 feet away from the central screen. A digital

[1] Excluded infants' receptive vocabulary size (*M* = 161, range 4 – 383) was comparable to that of the infants whose data were included (see Table 1).

video camera was mounted flush with the lower edge of the 60” central screen so that infants’ visual attention could be monitored and recorded to a computer hard drive when seated on their parents lap. An experimenter administered the tasks from outside the room via a PC computer running custom software. The experimenter was blind to the experimental stimuli and monitored infant visual attention via the digital video feed.

Statistical Learning Task Materials and Procedure. The materials and procedure for the statistical learning task were the auditory stimuli used in Experiment 2 of Graf Estes et al. (2007). Specifically, during the **Auditory Familiarization phase**, infants listened to a 5.5 minute stream of concatenated consonant-vowel syllables while playing quietly on the floor of the sound-attenuated room. The familiarization stream was composed of 8 unique syllables that were combined into 4 statistically-defined disyllabic "words". There were two counterbalanced languages; Language 1 and Language 2. Language 1 words were *timay*, *dobu*, *gapi*, *moku* and Language 2 words were *pimo, kuga, buti, maydo*. The materials were generated by a trained female speaker who recorded sets of three-syllable sequences. The middle syllables from these sequences were spliced into a fluent speech stream (e.g., the sequences *timaydo* and *maydobu* were spliced to form the sequence *maydo*). This technique preserved appropriate coarticulation contexts for the target syllables, both within and across words, but resulted in no pauses or other reliable acoustic cues to word boundaries.

In each language, two of the words occurred with relatively high frequency. For example, in Language 1 *timay* and *dobu* occurred 180 times while *gapi* and *moku* occurred 90 times. The high frequency words occurred together in sequence 90 times (e.g., *timaydobu*; *dobutimay*), creating "partword" sequences that recurred across word boundaries 90 times (e.g., *maydo*; *buti*). At **Test,** we used these frequently-occurring partwords, as well as the low-frequency words (e.g., *gapi* and *moku),* which occurred with equal frequency. Importantly, the test part-words had internal TPs of .5, but the test words had internal TPs of 1.0.

Infants' listening preferences for low-frequency words vs. partwords was assessed using the HTPP. On each test trial a word or partword was repeated up to 25 times. For example, on the *gapi* test trial, 4 different tokens of *gapi* were repeated in a random order separated by .5 sec pauses. Each test trial was presented twice, for a total of 8 test trials. All infants heard the same set of test trials, but the sequences that were words for an individual infant familiarized with Language 1 were partwords for infants familiarized with Language 2, and vice versa. This allowed us to control for arbitrary listening preferences.

During the test phase, the experimenter used the custom software to control the presentation of auditory stimuli and record infants' listening times. The infant was seated on a parent's lap, and each test trial began with a flashing red circle displayed on a monitor located in front of the infant. Once the infant focused on the screen, a GIF animation of a steam train appeared on a monitor 90 degrees to the right or left, and the center circle disappeared. When the infant turned his or her head at least 30 degrees toward the side stimulus, the experimenter signaled the computer to present an auditory stimulus. The auditory stimulus continued to play until the infant looked away for at least two consecutive seconds, or until 35 seconds had elapsed. To train infants on the contingency between their looking behavior and the presentation of auditory stimuli, the test phase began with two trials that consisted of music.

Lexical-Processing Efficiency (LPE) Task Materials and Procedure. After infants completed the statistical learning task they were given a short break (if needed) before we assessed their LPE. The LPE task tested infants' ability to find the referents of several common English words (see Fernald et al., 2008 for an overview). On each trial, two realistic color pictures of familiar animals or objects appeared in the lower right and left corners of the central monitor. After 3 seconds, infants heard the phrase "Find the [target word]" and the trial continued for approximately 3 seconds. The phrases were naturally produced in an infant-directed register by a female speaker.

The target words were *kitty, doggie, birdie, baby, car,* and *shoe*. Infants were tested on each target word 4 times, with side of target presentation counterbalanced. Each picture occurred equally often as a target and a foil. After every 4th trial, infants viewed a reward trial consisting of a colorful image moving across the screen paired with encouraging remarks, such as "Good job!" and "Way to go!". Infants' looking behavior during this task was recorded directly to a hard drive for coding and analysis.

Vocabulary Size Assessment. After the experiment, parents completed the MacArthur-Bates Communicative Development Inventory (MCDI): Words and Gestures. This form contains a set of lexical items that are typically learned within the first 3 years, and we used a count of the words that parents indicated that their child understands (receptive vocabulary).

Results and Discussion

Data Preparation. Because the minimum time needed to hear a word or part-word sequence was approximately 1 second, trials on which infants listened for less than a second were not included. Infants who did not contribute at least 2 usable test trials of each type (word and partword) were excluded from analysis. We used a proportional preference measure to capture performance on the statistical learning task. A preference score for each infant was created by dividing her average listening time to word trials by the sum of her average listening time to word plus partword trials. Thus, scores above .5 indicate a stronger preference for words over partwords, and the higher the number the greater their word preference. The proportional preference measure normalizes across differences in overall listening time[2]. While both words and partwords had

[2] The results were largely the same when we used a subtraction measure of listening time to words minus part-words.

been heard equally often during familiarization, the words contained more reliable statistical structure, and can therefore be thought of as relatively “familiar”. The partwords, because they consisted of less reliable transitions, were relatively “novel”.

The recordings of the LPE task were coded frame-by-frame offline by trained observers naïve to the content of each trial using iCoder software developed by Fernald et al, (2008). Data from 10% of participants were randomly selected to be independently coded. Agreement between coders within a single frame was greater than 98%. Following Fernald and colleagues (2008), LPE was assessed on trials during which infants were looking to the distractor picture at label onset. In particular, the time it took to initiate a gaze shift to the correct picture was averaged across those trials to obtain a reaction time (RT) score. To increase the likelihood that shifts were related to hearing the target word, trials on which infants shifted before 367ms had elapsed, the minimum time needed to plan and initiate an eye-movement in response to auditory information, or after 2200ms, were not included. Using this task, Fernald has found that correlational analyses between RT calculated over all 6 words presented and MCDI scores do not yield different results from those using RT calculated using only words that parents reported their infant to understand (Fernald et al., 2006). Calculating RT over more trials is likely to provide a more stable estimate, and thus, we included trials testing all 6 words. Data from infants who did not contribute at least 2 usable trials over which to calculate a mean RT score were excluded from analysis.

Preliminary Analyses. All tests in this experiment were 2-tailed and alpha levels were set to .05. Preliminary analyses revealed no effects of artificial language version on infants' preferences for words, and no effects of sex for this measure or for the LPE and vocabulary size measures. Thus these factors were not included in further analyses. A one-sample *t*-test indicated that infants as a whole did not show a preference for either test trial type ($M = .5$, $SE = .02$, $t\ (37) = .11$, $p = .92$). We will address the implications of this finding in the context of the results of the correlational analyses.

Table 1: Experiment 1 Descriptive Statistics

	Mean	Min	Max	SD
Words Understood (raw)	154.4	37	375	95.2
Words Understood (percentile)	45.06	5	99	31.7
SL Task Preference	.50	.30	.72	.11
LPE Task RT (ms)	1003.8	533.7	1379.8	229.9

Note: The Words Understood raw and percentile scores came from the MCDI, the SL (Statistical Learning) Task Preference score reflects their preference for listening to words over partwords (a proportion score), and the LPE Task RT reflects the speed with which they used English words to find target pictures.

Relations Between Performance on the Statistical Learning and LPE Tasks. Our primary question was whether individual differences in statistical learning are related to LPE. We addressed this question by testing how these factors covary. We found that they are negatively correlated, such that infants who showed stronger familiarity preferences had shorter RTs on the LPE task (See Table 2; Figure 1). Infants' preferences can change over the course of testing, and we found that the relation between preferences and LPE was stronger for the first test block (*r* (36) = -.33, *p* = .04) than in the second block (*r* (36) = -.25, *p* = .14). However, an ANCOVA indicated that there was no interaction between these factors (*F* (1, 34) = .47, *p* = .5), and thus, we did not break down the preference score by block in subsequent analyses.

Table 2: Experiment 1 Correlations

	Age	Vocabulary Size	LPE
Age	–	–	
Vocabulary Size	-.01	–	–
LPE	.05	-.33*	–
SL Task Preference	-.05	.16	-0.38*

* p <.05

Note: Infants' receptive vocabulary size (raw) was measured using the MCDI, SL (Statistical Learning) Task Preference by the proportion score reflecting their preference for words over partwords, and the LPE by the speed with which they used spoken words to find target pictures.

Because previous work suggests that infants' LPE scores are related to age and vocabulary size (Fernald et al, 1998; 2006), we also wanted to account for variance in LPE that was related these factors. To that end, we performed a hierarchical regression in which we tested whether statistical learning scores account for unique variance in LPE. In the Control Model we entered age and vocabulary size as control variables, and performance on the statistical learning task was entered in the next block of the analysis to test the SL model. The dependent variable was performance on the LPE Task. The Control model including age and vocabulary size did not predict variance in SPE; R^2 = .11, *F* (2, 35) = 2.15, *p* = .13 (see Table 3). Critically, adding statistical learning to the model (the SL Model in Table 3) accounted for significant additional variance in LPE ΔR^2 = .12, ΔF (1,34) = 5.02, p = .03 (see also Figure 1).

Table 3: Regression Models

Step 1 (Control Model)	***B***	***SE B***	**β**	***p***
Constant	533.16	1873.55		.76
Age	36.15	120.89	.05	.77
Vocabulary Size	-.79	.39	-.33	.05

Step 2 (SL Model)	***b***	***SE B***	**β**	***p***
Constant	1080.75	1789.21		.55
Age	24.96	114.60	.03	.90
Vocabulary Size	-.66	.37	-.27	.08
Statistical Learning	-725.82	323.94	-.33	.03

These findings suggest that infants' performance on a test of their ability to learn statistical structure in speech was selectively related to their native language LPE. Specifically, infants with stronger preferences for word test items were faster on the LPE task, but neither preference scores nor LPE were related to age. Furthermore, infants' vocabulary size predicted their LPE but performance on the statistical learning task accounted for additional variance.

As can be seen in Figure 1, infants who showed a familiarity preference tended to have faster LPE scores, and those with a novelty preference tended to have slower LPE scores. However, when the same artificial language was used to test TP learning in younger infants using a central fixation auditory preference procedure (Graf Estes, 2012), they showed a novelty preference at test. It is beyond the scope of the current work to determine why our findings differed from those of Graf Estes (2012) (i.e., whether the difference is related to the specific method we used, to the age of our participants, etc.). However, given that we did not predict a direction of preference a priori, we note the possibility of a type 1 error, or that we obtained this correlation by chance. Thus, Experiment 2 was designed as a conceptual replication in which we asked if sequential statistical learning and LPE are related when using different artificial-language materials that incorporate sequential statistics relevant to learning word-order patterns.

Experiment 2

The goal of Experiment 2 was to test whether infants' ability to learn sequential statistical regularities across words is related to their LPE. To that end, we presented 15-month-olds with an artificial language containing *aX bY* word-order relationships. This language incorporates sequential structure modeled after determiner-noun and auxiliary-verb co-occurrence relationships. Gómez and Lakusta (2004) found that when 12-month-olds were familiarized to a similar *aX bY* language, and then tested on grammatical and ungrammatical strings, most infants showed a preference for the grammatical strings. Thus, in the current experiment we used the HTPP to assess infants' ability to learn statistical regularities in the *aX bY* language, and assessed LPE as in Experiment 1. Based on the findings of Experiment 1 and the prior work with this artificial language, we predicted that infants showing a stronger preference for statistically robust structure (i.e., a stronger preference for familiar grammatical strings over ungrammatical ones) would also have better LPE.

Methods

Participants. Participants were 30 infants between 15.0 and 15.9 months of age (16 female; see Table 3 for additional descriptive statistics). The inclusion criteria were the same as in Experiment 1. An additional 14 infants were tested but their data were excluded for fussiness (*n*=8), equipment failure and experimenter error (*n*=4), or failure to contribute sufficient usable data (*n*=2).[3]

Materials and Procedure. The entire experiment took place in the sound-attenuated booth from Experiment 1, and the general structure was parallel, such that infants were trained and tested on an artificial language-learning task, then tested on the LPE task, and finally caregivers filled out the MCDI: Words and Gestures form.

Statistical Learning Task. During the Familiarization phase, infants were seated on a parent's lap as the artificial language was played. There were two *a*-words (*ong*, *erd*) and two *b*-words (*alt*, *ush*), and eight each of the *X*- and *Y*-words. The *X*-words were disyllabic (*coomo*, *fengle*, *kicey*, *loga, paylig, wazil, bevit, meeper*), while the *Y*-words were monosyllabic (*deech*, *ghope, jic*, *skige, vabe, tam, vot, rud*). Infants were familiarized to one of two versions of the language. In Language 1 strings took the form *aX* and *bY* (e.g., *erd coomo*, *ong kicey*, *ush*

[3] On average, excluded participants' vocabulary size (*M* = 129, range 22-338), which was comparable to that of the infants included in the final data set (See Table 3).

deech, *alt skige*). In Language 2 the pairings were switched such that strings took the form *aY* and *bX* (e.g. *erd deech*, *ong skige*, *ush coomo*, *alt kicey*).

A key feature of this artificial language is that the *X* and *Y* categories were distinguished by correlated phonological and distributional cues. Using Language 1 to illustrate, disyllabic *X*-words followed *a*-words in phrases (e.g., *ong coomo*, *erd coomo*), whereas monosyllabic *Y*-words followed *b*-words (e.g., *alt deech*, *ush deech*). Thus, words that shared phonological properties also shared distributional properties. The presence of correlated cues is critical to grouping words into categories, and to learning their co-occurrence relationships in this *aX bY* language (Gómez & Lakusta, 2004; Lany & Gómez, 2008).

The materials were spoken by an adult female in an animated voice and digitized for editing. One token of each word was selected so that *aX* and *bY* phrases (or *aY* and *bX* phrases in Language 2) could be created by splicing tokens together separated by 0.1-sec pauses. Thus, each language contained 32 unique phrases (16 *aX* and 16 *bY*, or vice versa in Language 2). Each phrase occurred four times during the randomized familiarization sequence, which lasted 3.5 minutes.

During Familiarization, infants were also trained on the contingency between their looking behavior and the appearance and disappearance of stimuli on the center and side monitors. When the infant was attentive, the experimenter used custom PC software to initiate the Familiarization stream and to control the presentation of the visual stimuli. At the same time, a flashing light was displayed on a monitor located in front of the infant. Once the infant was focused on the screen, the experimenter signaled the computer to present the steam train GIF on a monitor 90 degrees to the right or left, and the flashing light disappeared. The side stimulus was displayed until the infant looked away for more than 2 consecutive seconds, or more than 45 seconds had elapsed. At that point the center stimulus reappeared and the process repeated. The auditory stream played continuously while the front and side stimuli appeared contingent on infants' looking behavior.

After Familiarization, infants were tested on their ability to discriminate between grammatical and ungrammatical strings. The test included four trials consisting of Language 1 strings (e.g., *erd coomo*, *ush deech*, *ong fengle*, *alt ghope*) and four trials consisting of Version 2 strings (e.g., *erd deech*, *ush coomo*, *ong ghope*, *alt fengle*). Thus, all of the words in the test trials were familiar to infants, but not all of the phrases conformed to the familiarization language. For infants familiarized to Language 1, the Language 1 strings were Grammatical, and the Language 2 strings were Ungrammatical, and vice versa for infants familiarized to Language 2. Critically, infants could only distinguish between Language 1 and 2 strings if they had learned the co-occurrence relationships between the words.

Infants were presented with two blocks of test trials, each consisting of two Language 1 and two Language 2 trials, for a total of 8 test trials. The custom software was used to control the presentation of the auditory stimuli and to record infants' listening times, as in Experiment 1. Trials lasted until the infant looked away for more than 2 consecutive seconds or more than 45 seconds elapsed. The order and side of test-trial presentation were randomized.

Results and Discussion

Data Processing. The HTPP data were processed similarly to Experiment 1, with one exception. Because strings were approximately 1.5 seconds in duration, trials on which infants' listening times were less than 1.5 second were not included in the calculation of their mean listening times. As in Experiment 1, infants who did not contribute at least 2 usable test trials of each type (Grammatical and Ungrammatical) were excluded from analysis. Infants' preference scores were calculated as in Experiment 1, as were their RTs on the LPE task. LPE data from 10% of participants were randomly selected and independently coded. Agreement between coders within a single frame was greater than 98%.

Preliminary Analyses. Infants' performance on the LPE and statistical-learning tasks did not differ as a function of sex or language version, and thus we collapsed across these variables in subsequent analyses. As in Experiment 1, infants did not show a significant preference for the familiar test items (M = .51, SE = .03; t (29) = .2, p = .4).

Relations Between Performance on the Statistical Learning and LPE Tasks. Our main question was whether infants' statistical learning ability was related to their LPE. We used a 1-tailed correlation for our initial analyses because we predicted that infants with greater familiarity preferences would also process native language speech more rapidly. Overall, infants' preferences were negatively correlated with their LPE, but the strength of the relation was quite weak, and did not reach significance; r (29) = -.13, p > .05. As in Experiment 1, the relation was significant in the first block of testing r (29) = -.31, p = .05, but not the second r (25) = .21, p > .05 (Table 5). A two-tailed ACOVA revealed an interaction between test block (Blocks 1 and 2) and LPE (F (1, 24) = 4.24, p = .05. Only Block 1 was related to SPE; B = -14.02, p = .03 (Block 2 B = 4.35, p = .39). Thus, we focused on Block 1 performance in subsequent analyses.

Table 4: Experiment 2 Descriptive Statistics

	Mean	Min	Max	SD
Words Understood (raw)	140.9	19	343	76.72
Words Understood (percentile)	40.32	3	97	19.65
SL Task Preference (Block 1)	.48	.09	.93	.24
LPE RT (ms)	1052.7	533.5	1933.3	320.5

Note: As in Experiment 1, the Words Understood raw and percentile scores came from the MCDI, the SL (Statistical Learning) Task score reflects their preference for listening to Grammatical over Ungrammatical strings (a proportion score), and LPE RT reflects the speed with which they interpreted familiar English words.

To determine whether LPE was selectively related to statistical learning performance, we used a hierarchical regression in which infants' age and receptive vocabulary size were entered in the first step, and statistical learning (Block 1) was entered in the second (see Table 4 for descriptives). The Control Model including age and vocabulary size did not explain significant variance in SPE; R^2 = .07, F (2, 27) = .95, p = .4 (see Table 6). However, the adding infants' statistical learning scores (the SL Model in Table 6) increased the model fit, ΔR^2 = .13, ΔF (1, 26) = 4.5, p = .04.[4]

Table 5: Experiment 2 Correlations

	Age	Vocabulary	LPE
Age	–	–	–
Vocabulary	0.12	–	–
LPE	-0.07	0.24	–
SL Task Preference	-0.38**	-0.01	-0.31*

Note: Receptive vocabulary size (raw) was measured using the MCDI, LPE by the speed with which they used English words to find target pictures, and SL (Statistical Learning) by their preference for listening to Grammatical over Ungrammatical strings in Block 1 (a proportion score).
* p <.1, **p <.05

Table 6: Regression Models

Step 1 (Control Model)	***B***	***SE B***	**β**	***p***
Constant	2853.41	3537.35		.81
Age	-126.00	229.67	-.10	-.55
Vocabulary Size	1.03	.78	.25	1.32

Step 2 (SL Model)	***b***	***SE B***	**β**	***p***
Constant	6018.74	3649.32		.11
Age	-314.98	233.85	-.26	.19
Vocabulary Size	1.09	.74	.26	.15

[4] As in Experiment 1, the pattern of findings across the analyses was unchanged when we used mean listening to Grammatical strings minus listening to Ungrammatical strings, rather than the proportional score used here.

Statistical Learning	-526.66	248.72	-.40	.04

In sum, we found that 15-month-old infants' LPE is related to their performance on a task assessing learning of word-level co-occurrence statistics in an artificial language. Specifically, infants' preference for familiar grammatical strings vs. ungrammatical strings predicted their LPE, such that infants who tended to listen longer to familiar, grammatical strings processed familiar native language words more rapidly. This relation, while holding only in the first block of testing, was significant when including variance related to infants' age and vocabulary size.

These findings parallel those from Experiment 1, in which infants with better LPE showed a preference for highly predictable syllable sequences. Thus, across two different tasks, two groups of infants showed the same pattern of performance. Infants who encoded native language speech more efficiently tended to listen longer to sequences that were consistent with the most robust statistical regularities in their familiarization, while infants with slower LPE scores tended to listen longer to sequences that were inconsistent with that structure. In both experiments we also found relations between native language processing and statistical learning when including variance accounted for by age and vocabulary size. Nonetheless, the relations we observed are modest, and should be taken as a starting point for future research.

General Discussion

We tested whether infants' sequential statistical learning ability is related to their ability to encode and interpret native-language speech. We used two classic artificial language structures that tap infants' learning of sequential structure: One tapped statistical regularities holding across syllables, or TPs (Experiment 1) and the other tapped sequential regularities holding across words (Experiment 2). Previous work suggested that infants aged 15 months and younger are capable of learning the kinds of sequential structure found in both artificial languages (Aslin et al., 1998; Gómez & Lakusta, 2004; Graf Estes et al., 2007; Graf Estes, 2012; Lany & Gómez, 2008). We assessed speech processing using a well-studied lexical-processing efficiency (LPE) task in which infants heard words that were likely to be familiar to them. The words were presented in simple ostensive labeling phrases that are common in speech to infants (Cameron Faulkner et al., 2003).

In Experiment 1, we found that infants who showed stronger preferences for word-like units from an artificial language also performed better on the LPE task. In Experiment 2, we found that infants who showed stronger preferences for attested word-order patterns also had better LPE, though infants' preferences faded over the course of testing. The relations we observed were of modest strength, but there was consistency in the results. Together, these findings suggest that infants' ability to learn novel sequential regularities is related to their ability to encode and interpret fluent speech from their native-language.

While we found evidence that sequential statistical learning abilities and LPE are related, we should note that we did not predict that better statistical learning would be evidenced as a novelty preference; in fact, we did not predict a specific direction of preference. Researchers often remain uncommitted about the likely the direction of preference that they will obtain in a given HTPP experiment, as both familiarity and novelty preferences are common. For example, Gómez and Lakusta (2005) found familiarity preferences using an *aX bY* language similar to the one used in Experiment 2, as we did, while other studies using artificial language materials have found novelty preferences (e.g., Marcus et al., 1999; Saffran et al., 1996). Some studies have also reported both familiarity and novelty preferences across related experiments as a function of ages and training or testing conditions (e.g., Gómez & Maye, 2005; Thiessen & Saffran, 2003).

Nonetheless, interpreting our HTPP results requires careful consideration. In two recent studies, stronger familiarity preferences on a segmentation task using natural language materials were associated with better native language development (Singh et al., 2012, Newman et al., 2006; but see Newman et al., 2016 for a different result). Our finding that stronger familiarity preferences for attested structure in artificial-language materials were associated with better LPE is consistent with these studies, and with an interpretation in which better statistical learning in related to better LPE. While it is possible that greater novelty preferences actually reflected better learning, we suggest that this is unlikely: Such an interpretation would mean that infants with better statistical learning skills were less proficient with their native language, but this pattern of findings is more characteristics in studies in which the novel statistical regularities are designed to *violate* native language phonotactics (e.g., Graf Estes et al., 2011; 2016). In such studies, the effects tend to become more pronounced with age. However, infants older than those we tested persist in tracking the statistical regularities in the artificial languages we used here (Graf Estes et al., 2007; Lany & Saffran, 2010), and better performance has also been associated with better language development (Lany, 2014). The interpretation that greater familiarity preferences reflect better statistical learning, and consequently that better statistical learning is associated with better LPE, is also consistent with evidence that the ability to learn statistical regularities is related to language processing skills in adults (Conway et al., 2010; Misyak & Christiansen, 2012) and school aged children (Kidd & Arciuli, 2016; Spencer et al., 2015).

In future work it will be important to continue to investigate how statistical learning is related to real-time language processing. Interpreting relations between HTPP tasks assessing statistical learning and other measures of language learning and development would be easier if group-level direction-of-preference effects were more straightforward and if it were possible to make a priori precise direction-of-preference predictions. In principle, the HTPP is well-positioned to capture both group effects and individual variability (Kemler Nelson et al., 1995), but the parameters of the task may need to be more extensively tailored to a given set of materials at a specific age or developmental level. Thus, it will be valuable to refine tests of infant sequential statistical learning in future work.

This issue notwithstanding, our data may have implications for studies using the HTPP that yield null results. A common approach taken to facilitate interpretation of direction of preference is to test infants in a narrow age-range, or to contrast performance across several narrow age ranges. The current data suggest that even if a group of infants at a specific age show no statistically-reliable preference overall, it may not mean that infants failed to distinguish between different types of trials (in this case statistically attested vs. unattested structure). Infants' age is good predictor of language learning development at a gross level (24-month-olds are generally more advanced than 18-month-olds, and 18-month-olds more advanced than 12-month-olds), but not necessarily at a month-to-month level. For example, vocabulary size is better than age at predicting LPE and when infants will begin to combine words in their own speech (Borovsky, Elman, & Fernald, 2012; Bates et al., 1988). Thus, one potentially promising approach in studies employing the HTPP, including efforts to replicate previous work, would be to match participant samples on factors other than age, including measures of language processing. Such an approach might yield more meaningful data, and facilitate replicating previous findings.

Another caveat is that the correlational design we used cannot support inferences about causation. Our goal was to study the naturally occurring covariation between statistical learning ability and speech processing as a means of supplementing the extensive body of tightly controlled experimental work on statistical learning. However, we cannot rule out the possibility that the observed relations were driven by a third underlying factor. For example, it is possible that infants who both processed speech efficiently and readily learned statistical regularities were simply more advanced across the board. We were able to partially address this limitation by building on previous findings that there is a relation between measures of infants' vocabulary size, and measures of their LPE (Fernald et al, 1998; 2006; Weisleder & Fernald, 2013). Specifically, we found that statistical learning ability is also related to LPE, above and beyond the variance related to vocabulary size, providing some evidence for specificity in the relation. Nonetheless, it is possible that more advanced language skills like phonological encoding drive superior performance in both tasks. Thus, it would be valuable to test whether sequential processing in both artificial- and native-language materials is related to the precision of phoneme representations, as well as other aspects of phonological development. For example, it would be informative to test the precision of infants' phoneme representations, and the speed with which they access them, in conjunction with sequential statistical learning and LPE, to determine whether these factors explain additional variance in the measures, or in the relation between them. It would also be valuable to test relations between sequential statistical learning and LPE tasks that are tailored to tap different kinds of sequential structure. While a small number of studies have assessed differences in young children's processing efficiency for predictive vs. unpredictive structures or contexts, these tasks have been used minimally, and not at all with the younger ages tested here (Borovsky et al., 2016; Mani & Huettig, 2012). Thus, testing slightly older infants might provide important insights into this question.

Considered together, we suggest that the findings from these two experiments provide a platform on which future research on connections between sequential statistical learning and language development can build. For example, a critical step in understanding the mechanisms of language learning will be to determine *how* speech-processing efficiency and statistical learning ability are related. In the Introduction we suggested that statistical learning may support online speech processing, and specifically that the ability to detect and use predictive structure in speech via statistical learning is what contributes to gains in the ability to encode and interpret speech in real time. However, it is also possible that speech-processing ability gates the input to statistical learning mechanisms. In fact, it is likely that encoding and statistical learning processes influence each other across development, perhaps becoming more interdependent as infants gain language experience. A potential approach to determining how these processes are related would entail examining these relationships as they unfold across development (see Arciuli & Torkildson, 2012; Kidd & Arciuli, 2016), including measures of phonological processing.

In sum, previous studies have linked statistical learning with language outcomes (e.g., Graf Estes et al., 2011; Lany, 2014), and the current work builds on such findings by demonstrating a link between statistical learning and a specific language skill – lexical-processing efficiency. These data suggest that LPE may be a good way to test relations between statistical learning ability and language development. Furthermore, consistent with findings from work with adults (Conway et al., 2010), they suggest that one way that good statistical learning skills may affect infants' language development is by promoting the ability to encode speech rapidly.

References

Arciuli, J., & von Kos Torkildsen, J. (2012). Advancing our understanding of the link between statistical learning and language acquisition: the need for longitudinal data. *Frontiers in Psychology, 3,* doi:10.3389/fpsyg.2012.00324.

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.

Bates, E., Bretherton, I., & Snyder, L. (1988). From First Words to Grammar: Individual Differences and Dissociable Mechanisms. Cambridge: Cambridge University Press.

Borovsky, A., Ellis, E.M., Evans, J.L., Elman, J.L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development, 87*, 1893-1908.

Borovsky, A., Elman, J.L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults (2012). *Journal of Experimetnal Child Psychology, 112,* 417-436.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa K., Cabrol, D., & Christophe, A. (2014). Function words constrain online recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development, 10,* 1-18.

Conway, C.M., Bauernschmidt, A., Huang, S. S., Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114,* 356-371.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers*, *28*, 125–127.

Ellis, E.A., Robledo Gonzalez, M., & Deak, G.O. (2013). Visual prediction in infancy: What is the association with later vocabulary? *Language Learning and Development, 10*, 36-50,

Farmer, T.A., Christiansen, M.H., & Monaghan, R. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences, 103*, 12203-12208.

Fenson, L., Pethick, S., Renda, C., & Cox, J.L. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21,* 95-116.

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., & Marchman, V.A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Children Development, 83,* 203-222.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*, 98-116.

Fernald, A., Zangl, R., Portillo, A., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I.A. Sekerina, E.M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.

Gerken, L. A., Wilson, R., & Lewis, W. (2005). Seventeen-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431-436.

Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction in 12-month-old infants. *Developmental Science*, *7*, 567–580.

Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7*, 183-206.

Graf Estes, K. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00447

Graf Estes, K., Edwards, J., & Saffran, J.R. (2011). Phonotactic constraints on infant word learning. *Infancy, 16,* 180-197.

Graf Estes, K., Gluck, S. C., & Grimm, K. J. (in press). Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology, 146*, 34-49.

doi: 10.1016/j.jecp.2016.01.012

Houston-Price, C., & Nakia, S. (2004). Distinguishing Novelty and Familiarity Effects in Infant Preference Procedures. *Infant and Child Development, 13*, 341-348.

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

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27,* 236-248.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language.* The MIT Press. Cambridge, MA.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32,* 402–420.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*, 630-645. doi: 10.1006/jmla.1994.1030

Kaufman, S.B., DeYoung, C.G., Gray, J.R., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*, 321-340.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18-and 24-month-old infants' use of function words to determine reference. *Child*

Development, 77, 325-338.

Kemler Nelson, D.G., Jusczk, P.W., Mandel, D.R., Myers, J., Turn, A., & Gerken. L.A. (1995). The Heat-Turn Preference Procedure for testing auditory perception. *Infant Behavior and Development, 18,* 111-116.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology, 48*, 171-184.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development, 87*, 184-193.

Lany, J., & Gómez, R.L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science, 19*, 1247-1252.

Lany, J., & Saffran, J.R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science, 14*, 1207-1219.

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infant acquisition of lexical categories. *Psychological Science*, *21*, 284-291.

Lany, J. (2014). Judging words by their covers and the company they keep: Probabilistic cues support word learning. *Child Development, 85,* 1727-1739.

Lany, J., & Saffran, J.R. (2013) Statistical learning and language acquisition in infancy. In Tager-Flusberg (Eds). *Comprehensive Developmental Neuroscience, Volume 3,* (pages 231-248).

Law II, F., & Edwards, J. R. (2014). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development, 4*, 331-355.

Lew-Williams, C. & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*, 193-197.

Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: a usage-based approach. *Journal of Child Language, 30,* 333-370.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 843-847.

Marchman, V.A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11,* F9-F16.

Marcus, G.F., Vijayan, S., Rao, S.B., & Vishton, P.M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80.

Mintz, T.H., Newport, E.L., & Bever, T.G., (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science, 26*, 393-424.

Misyak, J.B., & Christiansen, M.H. (2012). Statistical learning and language: An individual differences study. *Language Learning: A Journal of Research in Language Studies, 62,* 302-331.

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

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology, 39*, 484-494. doi: 10.1037/0012-1649.39.3.484

Shafto, C. L., Conway, C.M., Field, S.L., & Houston, D.M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy, 17*, 247-271.

Singh, L., Reznik, J.S., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science, 15,* 482-495.

Spencer, M., Kaschak, M.P., Jones, J.L., & Lonigan, C., J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing, 28,* 467-490.

Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word-recognition. *Developmental Science, 12,* 419-425.

Waxman, S.R., Lidz, J.L., Braun, I.E., & Lavin, T. (2009). Twenty-four-month-old infants' interpretation of novel verbs and nouns in dynamic scenes. *Cognitive Psychology, 59*, 67-95.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological*

Science, 24, 2143-2152.

Figure Captions

Figure 1: Infants preference scores are calculated as the proportion of time spent listening to syllable sequences that had high TPs divided by total listening time.

Figure 2: Infants preference scores are calculated as the proportion of time spent listening to attested word-order patterns divided by total listening time.

Figure 1

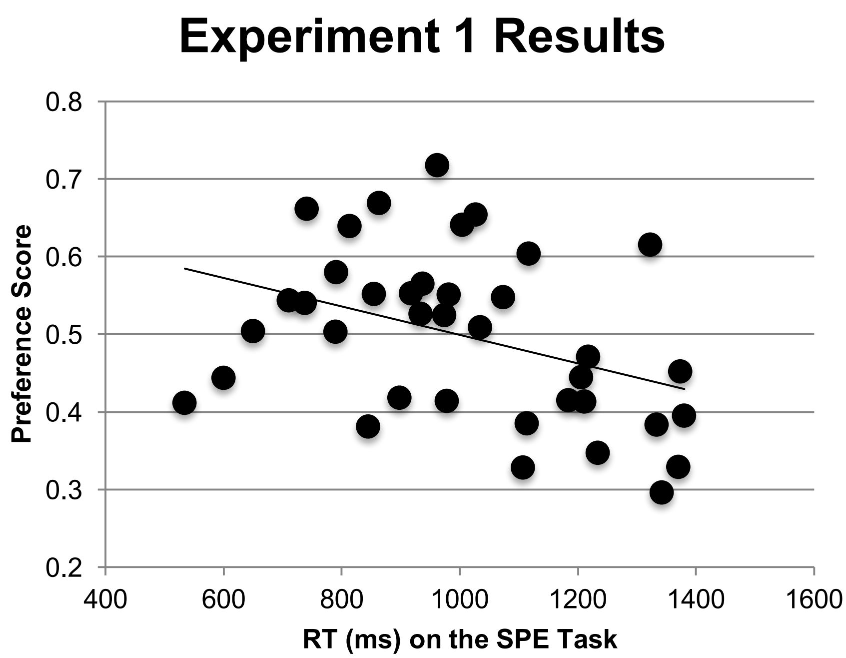
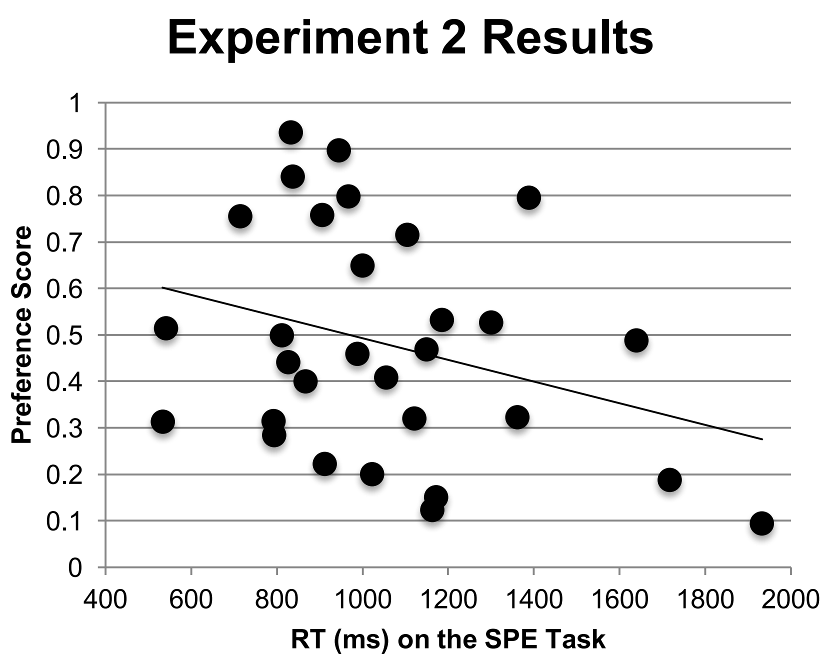

Experiment 1 Results
Preference Score
0.8
0.7
0.6
0.5
0.4
0.3
0.2
400
600
800
1000
1200
1400
1600
RT (ms) on the SPE Task

Figure 2:


Experiment 2 Results
Preference Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
400 600 800 1000 1200 1400 1600 1800 2000
RT (ms) on the SPE Task