# The learnability of Auditory Center-embedded Recursion

**Jun Lai (J.Lai@tilburguniversity.edu)**
Tilburg Center for Logic, Ethics and Philosophy of Science; Tilburg Center for Cognition and Communication, Tilburg University, the Netherlands

**Emiel Krahmer (E.J.Krahmer@tilburguniversity.edu)**
Tilburg Center for Cognition and Communication, Tilburg University, the Netherlands

**Jan Sprenger (j.sprenger@tilburguniversity.edu)**
Tilburg Center for Logic, Ethics and Philosophy of Science, Tilburg University, the Netherlands

**Abstract**

A growing body of research investigates how humans learn complex hierarchical structures with center-embedded recursion (Bahlmann, Schubotz, & Friderici, 2008; Poletiek & Lai, 2012). Increasing evidence indicates that properties of the learning input have an impact on learning this type of recursion. For instance, recent studies found that staged input, fewer unique exemplars and unequal repetition facilitate learning (e.g. Lai, Krahmer, & Sprenger, 2014; Lai & Poletiek, 2011, 2013). Most of these studies investigated learning center-embedded recursion through *visual* input, whereas few studies examined the processing of *auditory* input. In the current study, we test: 1) whether participants are able to learn center-embedded recursive structure from exclusively auditory input; 2) whether the facilitative cues (ordering and frequency distribution) are attuned to the auditory modality. Our results successfully demonstrate the learning of auditory sequences with center-embedded recursion, and replicated the effect with visual input in the previous study (Lai et al., 2014).

**Keywords:** auditory learning; artificial language; recursion; starting small; frequency distribution

## Introduction

During infancy, human beings start to demonstrate amazing abilities of obtaining useful information from numerous streams of auditory signals, which appear unsystematically in the environment. The crucial abilities enable humans to encode relevant information in a temporal order, since we do not receive all information at once (Conway & Christiansen, 2005). For instance, when we listen to an utterance, it is impossible to hear the whole sentence. Instead, we hear word by word. In order to process all the information, we first need to understand the relationship between segments. Statistical learning is a method to extract internal regularities or structural patterns from complex input (Romberg & Saffran, 2010).

Many statistical learning studies adopt the artificial grammar learning paradigm (Reber, 1967), which allows for investigating specific factors that affect language learning. It is a powerful tool to examine the cognitive mechanism of detecting statistical regularities from sequences that do not have a real-world meaning.

A large number of artificial grammar learning studies have shown that statistical learning mechanisms contribute to various aspects of language learning, such as word recognition, speech segmentation, etc. (Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). For example, empirical research that habituated infants to speech sequences following a statistical pattern found that 8-month-old infants were able to discover the pattern based on transitional probabilities between adjacent elements (Saffran, Aslin, & Newport, 1996). The probabilistic information of linguistic structures, such as frequency of occurrence, distribution of prosodic cues, or phonological patterns, can help learners detect regularities and improve learning (Romberg & Saffran, 2010).

Although studies have shown that simple grammar learning benefited from statistical information (Pena, Bonatti, Nespor, & Mehler, 2002), its role remains unclear in processing a higher level of grammar, for example, center-embedded recursive grammar (Mueller, Bahlmann, & Friederici, 2010). For example, "*The student that the teacher helped improved.*" is a typical center-embedded sentence. Due to the long distance dependencies between related elements, structures with center-embedded recursion are difficult to process and understand, but they are crucial in human language. Center-embedded recursion has been proposed to be a unique structure of human language (Fitch, Hauser, & Chomsky, 2005; Hauser, Chomsky, & Fitch, 2002).

Previous artificial grammar learning research, which tried to demonstrate the learnability of center-embedded recursion, provided diverging findings. Some studies observed various factors that enhanced learning, such as the staged input facilitation (Elman, 1993; Kersten & Earles, 2001). In a recent study, Lai and Poletiek (2011) trained participants with visual syllable sequences, generated by a hierarchical structured grammar, with the type of $A^{(n)}B^{(n)}$. Participants trained with staged input were compared with those trained with a random ordered input. For the staged input group, participants saw artificial grammar learning sequences in a "starting small" (SS) method, which arranged the input by increasing complexity. Thus, gradually, participants saw basic pairs with zero level of embedding

(0-LoE) first, then with one embedding (1-LoE), and two embedding (2-LoE) in the end. In the following classification test, participants were required to judge whether test items conformed to the same rule, which governed the previous learning input. Results showed that only the SS group was able to learn successfully and it outperformed the random group significantly. The finding of staged input effect was supported by a follow-up study (Lai & Poletiek, 2013), using a more complex form of staged input. By contrast, other studies did not observe any facilitation effect of incremental input (Fletcher, Maybery, & Bennett, 2000; Rohde & Plaut, 1999; Rohde & Plaut, 2003).

Another facilitative factor is skewed frequency distribution of input. In two experiments with visual center-embedded sequences, Lai and Poletiek (2013) found that learning was advanced, when the input was distributed unequally, favoring a larger number of basic exemplars (i.e. sufficient 0-LoE learning exemplars, fewer 1-LoE ones, and even fewer 2-LoE ones). The frequency distribution effect has also been found in other aspects, such as learning of verbs and phrases (Casenhiser & Goldberg, 2005; Kidd, Lieven, & Tomasello, 2010), long distance association in the structures such as A*X*B (Gomez, 2002), and grammatical categorization (Mintz, 2003).

Controlling for the frequency of various levels of embedding in the training set, Lai, Krahmer, and Sprenger (2014) investigated how the relative frequency of the learning exemplars would affect learning center-embedded recursion. They found that the diversity of various exemplars was not a necessity for successful learning of visual center-embedded sequences. Instead, training of fewer unique exemplars, but with repetition, could also lead participants to discover the complex recursive rule. Moreover, the more high-frequency exemplars occurred, the better participants learned. However, there are surprisingly few studies on facilitative cues, which aid in learning center-embedded recursion in the auditory modality. It deserves more attention in the field of artificial language learning, for a number of reasons. Firstly, at the initial stage of life, children learn a language first and foremost via the auditory modality. Empirical studies with infants have also stressed the importance of positive auditory experiences in early brain maturation (McMahon, Wintermark, & Lahav, 2012). The developed auditory modality helps children with information processing, language learning and memory formation (Moon & Fifer, 2000).

Secondly, modality has an impact on the performance of learning tasks (Huestegge & Hazeltine, 2011), and the sequential- or temporal way of presenting the input substantially determines the learning output (Conway & Christiansen, 2005). As shown in previous research on early brain development, children often learn their native language through the auditory modality, and refine their knowledge through the visual modality at a later stage (Holcomb & Neville, 1990). As regards to the modality difference, Glenberg and Fernandez (1988) found that the manner of temporal coding, in terms of the order of presentation, was more beneficial towards the auditory modality, compared to the visual modality, which relied more on spatial senses. Moreover, the greater variability in the auditory stimuli assists people in processing information. For example, patterns and regulations in rhythm (Rubinstein & Gruenberg, 1971) or in pitch (Evans & Treisman, 2010) of the input yielded learning differences, favoring the auditory modality but not the visual one. The statistical cues helped people detect auditory patterns in a more efficient way.

Thirdly, with regard to the staged input effect, Conway, Ellefson, and Christiansen(2003) compared a starting small group with a random group under both modalities. In Experiment 1 with visual letters, the starting small group was trained with increasing complexity (e.g. CW, CPTW, CPQMTW), whereas the random group received the same training material in a random order. In Experiment 2 with auditory material, the same input was adopted by replacing letters with consonant-vowel-consonant syllables. Conway et al. found a starting small effect for visual center-embedded structures, but not for auditory ones. They suggested that the lack of SS effect was due to intrinsic constraints of the auditory modality itself, since the auditory material appears in a temporal order. Note, however, that also Lai et al. (2014) presented the (visual) learning input syllable by syllable, emulating the auditory modality.

Last but not the least, studies have shown that the probability distribution of acoustics helped participants in speech perception (Clayards, Tanenhaus, Aslin, & Jacobs, 2008). This resembles the frequency effect found by Lai et al. (2014), but auditory stimuli were English words, instead of center-embedded recursive structures. To our knowledge, no previous research has probed into the frequency distribution effect in processing auditory center-embedded recursion.

This paper replaces visual stimuli with auditory ones and tests two main hypotheses: 1) whether humans can learn center-embedded recursion at all in the auditory modality, and 2) whether the facilitative cues (the ordering cue and the frequency distribution cue) are attuned to the auditory modality. We test participants' understanding and processing of the same set of center-embedded structures, but vary the training set. We compare learning performance under three conditions, i.e. Starting-small (SS), Starting-less (SL), and Starting-high (SH), copying the design of Lai et al. (2014). All conditions have the same number of training items (144) but differ in content. The SS condition provides an equal number of learning exemplars for each level of complexity (0-, 1-, 2-LoE). By presenting the input incrementally from the basic pairs to the most complex ones, learning difficulty is increased gradually. Compared to the SS group, the SL group has fewer unique exemplars

(36), which are repeated for an equal number of times (four times each). The SH group has also 36 unique exemplars, which are repeated unequally, depending on the exemplars' frequencies, which are skewed. For example, the number of occurrences of an item is higher if this is a high-frequent item. Thus, high-frequent items appear more often than low-frequent items. The SL and SH group both presented the input in a staged manner, according to the increasing complexity of exemplars.

## Experiment

### Method

**Participants.** Seventy-five students (54 female, mean age 21 year, SD 2.4) from Tilburg University participated for course credit[1]. All were native Dutch speakers. Participants had no prior knowledge about the experiment.

**Materials and design**. We applied the same set of syllable sequences as in Lai et al. (2014), which applied a grammar with the type of $A_nB_n$ and generated non-sense syllable sequences accordingly. A-syllables were [*be, bi, de, di, ge, gi*] and B-syllables were [*po, pu, to, tu, ko, ku*]. Each A-syllable was associated with a B-syllable according to its consonant pair. For example, *be/bi* was related with *po/pu*, *de/di* with *to/tu*, and *ge/gi* with *ko/ku*. Sequences consist of two, four, or six syllables (e.g. *bipo*, *bebepopo*, *gebiditopoku*). A Dutch speaker recorded all sequences. Reading speed, pitch and intonation were held constant. The recording time for each syllable was around 400 ms.

The test set, which was the same for all groups, consisted of 72 sequences, half grammatical and half not. The number of sequences for each level of complexity (i.e. 0-, 1-, and 2-LoE) was equal. Ungrammatical sequences were formed by mismatching an A-syllable with an unrelated B-syllable (i.e. be***ku***).

**Procedure.** Participants were randomly assigned to one of the three groups, 25 each. In the training phase, participants were required to attentively listen to sequences of sounds. The instruction stated that there was a rule underlying the sounds that they heard. Every trial began with a beep, followed by a sequence of sound, such as *bebepopo*. Each sound was displayed individually. In the test phase, participants were informed that they would hear new sounds, some of which obeyed the same rule as that in the previous training set, while some did not. Their task was to judge which test sounds followed the same rule. No feedback on answers was given during the test.

The whole experiment took approximately 30 minutes.

### Results

Figure 1(a) depicts the individual accuracies. Figure 1(b), which shows the group mean, indicates a similar learning pattern across conditions in both auditory modality and visual modality (Lai et al., 2014). A one-sample t-test showed that all groups achieved above chance performance significantly: $M_{SS}$= .55, $SE_{SS}$=.01, $t$ (23) = 3.13, $p$ = .005, $r^2$ = .30; $M_{SL}$= .57, $SE_{SL}$= .01, $t$ (24) = 4.54, $p$ < .001, $r^2$ = .46; $M_{SH}$= .62, $SE_{SH}$=.02, $t$ (23) = 7.86, $p$ < .001, $r^2$ = .73. The results suggested that these three groups succeeded in classifying grammatical test sequences from ungrammatical ones, to different extent.

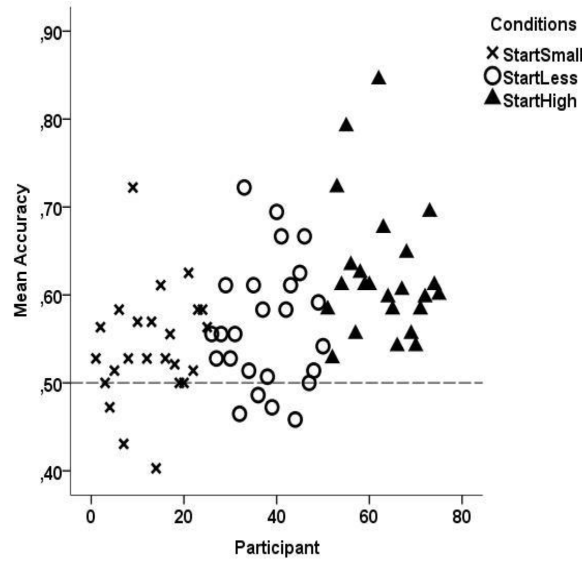


Figure 1(a). Scatterplot of individual accuracy. The dotted line represents chance level (M= .50).

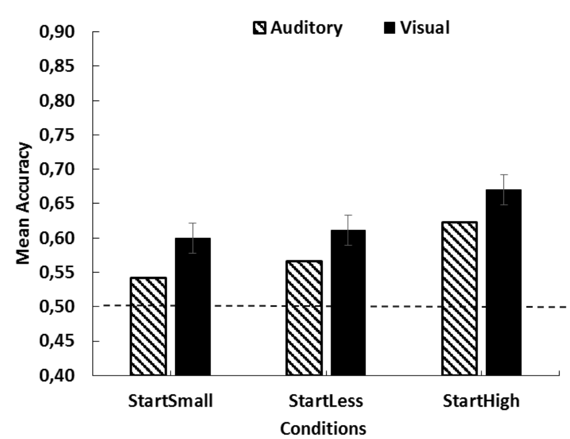


Figure 1(b). Mean accuracy of all conditions in both auditory and visual modality (Lai et al., 2014). The dotted line represents chance level (M= .50). Error bars indicate standard error of the mean.

We conducted a repeated-measure analysis, with Condition as the between-subjects factor, Grammaticality and LoE as within-subjects factors. The analysis first

[1] Two of these 75 participants were excluded from the data analysis due to interrupted termination of the experiment.

indicated a main effect of Condition, $F$ (2, 70) = 8.14, $p$ = .001, $\eta_p^2$ = .189. A post hoc Bonferroni test revealed that the SH group surpassed the SS group ($p$ =.001) and the SL group ($p$ =.021) significantly, while no significant difference between the SS and the SL group ($p$ = .715) was observed.

In addition, we conducted a dprime calculation, which was consistent with the calculation on mean accuracy. It also demonstrated a main effect of condition: $F$ (2, 70) = 7.95, $p$ = .001, $\eta_p^2$ = .185. The dprime scores were: $d'_{SS}$= .22, $SE_{SS}$=.34, $d'_{SL}$= .35, $SE_{SL}$=.39, $d'_{SH}$= .65, $SE_{SH}$=.45.

The analysis further showed a main effect of Grammaticality, $F$ (1, 70) = 5.95, $p$ = .017, $\eta_p^2$ = .078. The general score on grammatical test sequences (M= .60, SE= .01) was significantly higher than that on ungrammatical ones (M= .55, SE= .01), $p$ = .017. Specifically, there was a main effect of Condition on ungrammatical sequences, $F$ (2, 70) = 6.06, $p$ = .004, $\eta_p^2$ = .147, but no effect on grammatical ones, $F$ (2, 70) = 2.20, $p$ = .119. On ungrammatical sequences only, the SH group (M= .62, SE= .02) outscored the SS group (M= .52, SE= .02), $p$ = .007, and the SL group (M= .53, SE= .02) significantly, $p$ = .016.

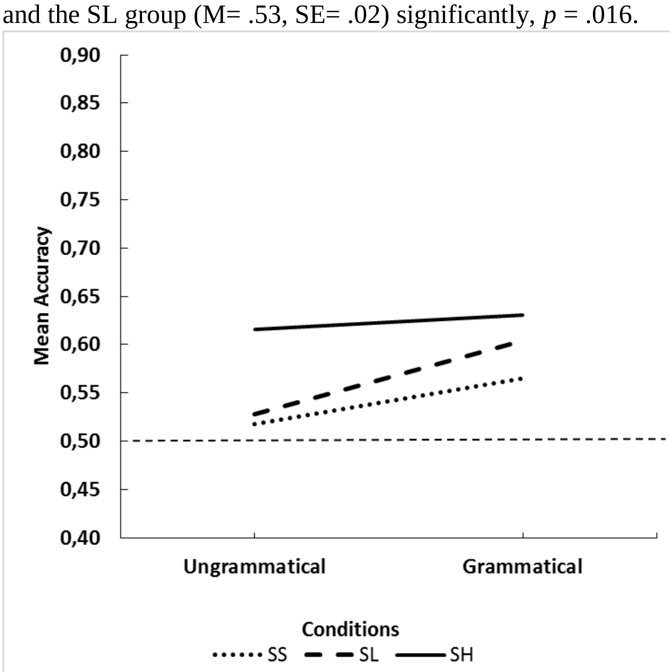


Figure 2. Mean accuracy of all conditions on ungrammatical and grammatical test sequences. The dotted line represents chance level (M= .50).

In order to pinpoint the substance of the facilitative effect, we examined the performance in different conditions at each level of complexity. For the SS group, only scores on 0-LoE (M= .61, SE= .02) were significantly above chance, $t$ (23) = 4.74, $p$ < .001, $r^2$ = .49. This indicated that the SS manner in the current study only helped participant make strong associations between the basic related pairs. However, for the SL group, performance on both 0-LoE (M= .64, SE= .03) and 1-LoE (M= .56, SE= 02) outperformed chance level, $t$ (24) = 4.93, $p$ < .001, $r^2$ = .50, and $t$ (24) = 3.42, $p$ = .002, $r^2$ = .33, respectively. Similarly, for the SH group, scores on 0-LoE (M= .77, SE= .03), $t$ (23) = 9.07, $p$ < .001, $r^2$ = .78, and those on 1-LoE (M= .58, SE= .03), $t$ (23) = 2.84, $p$ = .009, $r^2$ = .05, were both significantly above chance level, while scores on 2-LoE (M= .53, SE= .02) did not differ from chance, $t$ (23) = 1.13, $p$ = .270.

## Discussion

In the current study, we investigated the learnability of center-embedded recursive structures in the auditory modality. We also examined whether the facilitative factors, which aided in learning visual center-embedded recursion, were also applicable for auditory stimuli. First, participants in the auditory modality achieved significantly better than chance performance, independent of the relevant facilitative cue. These results markedly differ from the previous findings by Conway et al. (2003). One possible explanation is that their study used consonant-vowel-consonant syllables, such as *"biff", "rud", "sig"*, etc. Examples of their auditory sequences were *"biff-nep"* (0-LoE), *"biff-vot-cav-nep"* (1-LoE), etc. There were no salient acoustic cues implanted in these sound sequences. Nevertheless, in the current design, there are inherent acoustic regularities underlying the sequences. The first regularity is that all A-syllables end with *–e/-i* and B-syllables end with *–o/-u*. The second pattern is that A-syllables were connected with B-syllables, depending on the consonant pairs. The presence of phonological information might assist our participants first to realize the categorization of A-/B-syllables, and then discover the relation between associated elements. Therefore, our results challenged the claim that the lack of learning center-embedded recursion through auditory input was due to the modality itself. Instead, it might be caused by lack of sufficient acoustic information indicating the statistical relationship.

Secondly, we observed all three types of facilitative cues, i.e. staged input (SS), fewer exemplars (SL), unequal frequencies (SH), advanced learning center-embedded recursions in the auditory modality. There was no significant difference between the SS and the SL group, but the SH group surpassed these two groups significantly. In our experiment, the traditional SS setting is demonstrated to be useful in processing auditory center-embedded recursion. Compared to the SS group, the other two groups obtained much fewer unique exemplars. This poverty in exemplar diversity did not hinder learning. Instead, it helped participants focus on the statistical properties of the relatively small set of samples. It also fits humans' cognitive processing window, which deals with segments of information more efficiently (Christiansen & MacDonald, 2009). Furthermore, the large amount of repetition of these unique exemplars not only familiarizes participants with the acquired knowledge, but also consolidates their memories during learning. This indicates that a large number of various exemplars might not be necessary for learning complex center-embedded structures, even in the auditory modality. Instead, a repetition of a smaller set of unique but representative exemplars accelerates learning. For the SH group, the number of repetitions were unequal for exemplars with different frequencies. This arrangement of unequal repetition boosted learning, since participants were highly familiar with the most probable and typical structure in the grammar. The discovery of the most fundamental pairs aids in unpacking the complex syntactical structures.

Thirdly, regarding to the grammaticality of test items, we found that for all groups (SS+SL+SH), the general score on grammatical test items was significantly higher than that on ungrammatical ones. As Vokey and Brooks (1992) pointed out, participants were likely to compare the test items with their memorized exemplars and make their judgments based on similarity. Although test items are novel, the grammatical ones follow the same underlying rule and possess higher similarity to the learning items. Ungrammatical items might have been harder to judge because of the absence of a similarity cue. Interestingly, both in visual and auditory modality, the groups did not differ much in judging grammatical test items, However, for ungrammatical test items in the auditory modality, the SH group was more accurate than the other groups. This result is in line with the finding of Lai et al. (2014) for the visual modality. A possible explanation is that the unequal number of repetition fits an efficient way of cognitive processing, by giving prominence to the most representative structures.

Lastly, in accordance with the previous study with visual input (Lai et al, 2014), our results revealed that when the complexity of auditory input increased, the accuracy of grammaticality judgment decreased. The only difference is that the study with visual input found the performance of the SH group on 0-, 1-, 2-LoE items were all significantly better than chance. However, with auditory input, the SS group only scored significantly better than chance on 0-LoE, whereas the SL and the SH group achieved better than chance performance on 0-, and 1-LoE, but not on 2-LoE. This suggests that the successful learning of these two groups was not merely due to the recognition of basic exemplars (0-LoE), but also due to accurate judgments of more complex structures with embedding (1-LoE), though the most complex ones (2-LoE) seem too difficult for learning within such a short exposure. The results indicated that with auditory stimuli, the SS regimen might only advance learning at the basic level, i.e. the fundamental associations (0-LoE). Nevertheless, the SL and the SH setting can promote learning to a higher level. Thus, it seems more demanding in the auditory modality than in the visual modality to process higher level of complexity in the recursive hierarchy (2-LoE). Since the previous study (Lai et al., 2014) also controlled for the manner how visual stimuli were presented, the temporal order of auditory stimuli is not the primary reason. As Conway and Christiansen (2006) suggested, statistical learning under these two modalities is driven by separate subsystems and is guided by different sensory mechanisms. Memory constraints and other cognitive loads might prohibit the processing of auditory long-distance dependencies.

## Conclusion

In the present study, we demonstrate for the first time that participants were able to learn center-embedded recursion in the auditory modality, with the assistance of staged input. Our results challenge the view that the modality constraints prevented learning center-embedded recursion through the auditory modality. Furthermore, we also observed the starting small (SS), starting less (SL) and starting high (SH) effect with auditory input: staged input and the repetition of a smaller set of unique exemplars can promote efficient learning. So does the unequal number of repetition according to exemplars' frequencies. The results of the current auditory study coincide with those of the previous visual study. One possible reason is that Lai et al. (2014) did not use the traditional method to present visual sequences as a whole (Conway et al., 2003; Reber, 1967). Instead, they presented the visual sequences in a temporal order, i.e. syllable-by-syllable, to simulate the sequential order of auditory stimuli.

Our findings shed light on how statistical information of the input contributes to learning complex syntactical structure in the auditory modality. We manipulated three factors, i.e. staged input, repetition of exemplars, and unequal distribution in the statistical learning task. These three manipulations highly resemble a child-directed speech environment, which contains a large amount of simple structures but fewer complex sentences. Especially, the utterances are constantly repeated, for an unequal number of times (Snow, 1972). Further testing is worthwhile to verify the validity of auditory facilitation effect in natural language learning.

## References

Bahlmann, J., Schubotz, R.I., & Friederici, A.D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage, 42(2)*, 525-534.

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. Developmental Science, 8(6), 500-508.

Christiansen, M. H., & MacDonald, M. C. (2009). A Usage-Based Approach to Recursion in Sentence Processing. Language Learning, 59, 126-161.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. Cognition, 108(3), 804-809.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. Journal of Experimental Psychology-Learning Memory and Cognition, 31(1), 24-39.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities - Pitting abstract against stimulus-specific representations. Psychological Science, 17(10), 905-912.
Conway, C. M., Ellefson, M. R., & Christiansen, M. H. (2003). When Less is Less and When Less is More: Starting Small with Staged Input. Paper presented at the Proceedings of the 25th Annual Conference of the Cognitive Science Society, Mahwah.
Elman, J. L. (1993). Learning and Development in Neural Networks - the Importance of Starting Small. Cognition, 48(1), 71-99.
Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. Journal of Vision, 10(1).
Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. Journal of experimental psychology. Learning, memory, and cognition, 28(3), 458-467.
Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. Cognition, 97(2), 179-210.
Fletcher, J., Maybery, M. T., & Bennett, S. (2000). Implicit learning differences: A question of developmental level? Journal of Experimental Psychology-Learning Memory and Cognition, 26(1), 246-252.
Glenberg, A. M., & Fernandez, A. (1988). Evidence for Auditory Temporal Distinctiveness - Modality Effects in Order and Frequency Judgments. Journal of Experimental Psychology-Learning Memory and Cognition, 14(4), 728-739.
Gomez, R. L. (2002). Variability and detection of invariant structure. Psychological Science, 13(5), 431-436.
Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? Science, 298(5598), 1569-1579.
Holcomb, P. J., & Neville, H. J. (1990). Auditory and Visual Semantic Priming in Lexical Decision - a Comparison Using Event-Related Brain Potentials. Language and Cognitive Processes, 5(4), 281-312.
Huestegge, L., & Hazeltine, E. (2011). Crossmodal action: modality matters. Psychological research, 75(6), 445-451.
Kersten, A. W., & Earles, J. L. (2001). Less really is more for adults learning a miniature artificial language. Journal of Memory and Language, 44(2), 250-273.
Kidd, E., Lieven, E. V. M., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: evidence from sentential complements. Language Sciences, 32(1), 132-142.
Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. Cognition, 83(2), B35-B42.
Lai, J., Krahmer, E. J., & Sprenger, J. M. (2014). Studying Frequency Effects in Learning Center-embedded Recursion. In M. Knauff, M. Pauen, N.Sebanz, & I. Wachsmuth (Eds.), Proceeding of the 35th Annual Conference of the Cognitive Science Society (pp. 797-802). Austin, TX: Conitive Science Society.
Lai, J., & Poletiek, F. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. Cognition, 118(2), 265-273.
Lai, J., & Poletiek, F. H. (2013). How "small" is "starting small" for learning hierarchical centre-embedded structures? Journal of Cognitive Psychology, 25(4), 423-435.
McMahon, E., Wintermark, P., & Lahav, A. (2012). Auditory brain development in premature infants: the importance of early experience. Neurosciences and Music Iv: Learning and Memory, 1252, 17-24.
Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. Cognition, 90(1), 91-117.
Moon, C. M., & Fifer, W. P. (2000). Evidence of transnatal auditory learning. J Perinatol, 20(8 Pt 2), S37-44.
Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2010). Learnability of Embedded Syntactic Structures Depends on Prosodic Cues. Cognitive Science, 34(2), 338-349.
Pena, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. Science, 298(5593), 604-607.
Reber, A. S. (1967). Implicit Learning of Artificial Grammars. Journal of Verbal Learning and Verbal Behavior, 6(6), 855-863.
Rohde, D. L. T., & Plaut, D. C. (1999). Language acquisition in the absence of explicit negative evidence: how important is starting small? Cognition, 72(1), 67-109.
Rohde, D. L. T., & Plaut, D. C. (2003). Less is less in language acquisition. In P. Quinlan (Ed.), Connectionist modelling of cognitive development. Hove, UK: Psychology Press.
Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. Wiley Interdisciplinary Reviews-Cognitive Science, 1(6), 906-914.
Rubinstein, L., & Gruenberg, E. M. (1971). Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. Perception & Psychophysics, 9, 385-390.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. Science, 274(5294), 1926-1928.
Vokey, J. R., & Brooks, L. R. (1992). Salience of Item Knowledge in Learning Artificial Grammars. Journal of Experimental Psychology-Learning Memory and Cognition, 18(2), 328-344.