COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2018) 1–16

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12692

# Statistical Learning Is Not Age-Invariant During Childhood: Performance Improves With Age Across Modality

Amir Shufaniya, Inbal Arnon

*Department of Psychology, The Hebrew University of Jerusalem*

Received 9 March 2018; received in revised form 17 August 2018; accepted 30 August 2018

**Abstract**

Humans are capable of extracting recurring patterns from their environment via statistical learning (SL), an ability thought to play an important role in language learning and learning more generally. While much work has examined statistical learning in infants and adults, less work has looked at the developmental trajectory of SL during childhood to see whether it is fully developed in infancy or improves with age, like many other cognitive abilities. A recent study showed modality-based differences in the effect of age during childhood: While visual SL improved with age, auditory SL did not. This finding was taken as evidence for modality-based differences in SL. However, since that study used auditory linguistic stimuli (syllables), the differential effect of age may have been driven by stimulus type (linguistic vs. non-linguistic) rather than modality. Here, we ask whether age will affect performance similarly in the two modalities when non-linguistic auditory stimuli are used (familiar sounds instead of syllables). We conduct a large-scale study of children's performance on visual and non-linguistic auditory SL during childhood (ages 5–12 years). The results show a similar effect of age in both modalities: Unlike previous findings, both visual and non-linguistic auditory SL improved with age. These findings highlight the stimuli-sensitive nature of SL and suggest that modality-based differences may be stimuli-dependent, and that age-invariance may be limited to linguistic stimuli.

*Keywords:* Statistical learning; Language learning; Visual; Auditory; Developmental trajectory; Age-invariance; Children

## 1. Introduction

In the past two decades, much research has shown that humans are capable of extracting distributional regularities from their environment (see Romberg & Saffran, 2010;

Correspondence should be sent to Inbal Arnon, The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 9190501, Israel. E-mail: inbal.arnon@mail.huji.ac.il

Thiessen & Erickson, 2015 for reviews). This ability, often called statistical learning (SL), is present from early infancy (Saffran, Aslin, & Newport, 1996), including in newborns (e.g., Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009). Statistical learning is found across different modalities (i.e., auditory, visual, and tactile; Conway & Christiansen, 2005; Emberson, Conway, & Christiansen, 2011), and it can be used to learn various linguistic relations, from phonetic categories (Maye, Werker, & Gerken, 2002), through word order (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008), to phrasal structure (Gómez & Gerken, 1999). In line with its postulated role in language acquisition, SL performance is correlated with language outcomes in both children (e.g., Arciuli & Simpson, 2011; Kidd, 2012) and adults (e.g., Conway, Bauernschmidt, Huang, & Pisoni, 2010; Misyak & Christiansen, 2012).

While much work has studied SL in infants and adults, fewer studies have examined SL in children or looked at its developmental trajectory during childhood. Such developmental findings are crucial for understanding whether SL is fully developed in infancy or whether it improves with age, like many other cognitive capacities. Statistical learning is a form of implicit learning (Perruchet & Pacton, 2006) which can be broadly defined as distributional learning that occurs without explicit awareness. While SL and implicit learning are often studied in separate literatures, they share many commonalities (see Christiansen, in press). Traditionally, implicit learning is claimed to be age-invariant (Meulemans, Van der Linden, & Perruchet, 1998; Reber, 1993), a claim supported by studies showing no difference in accuracy between children and adults (Bertels, Boursain, Destrebecqz, & Gaillard, 2015; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, very few cognitive skills do not improve with age, and indeed there is growing evidence within the implicit learning literature that performance improves during childhood (e.g., Janacsek, Fiser, & Nemeth, 2012; Lukács & Kemény, 2015). SL may show similar developmental changes when assessed using cross-sectional samples of sufficient size and over larger age ranges (Arciuli & von Koss Torkildsen, 2012). In line with this prediction, a large-scale cross-sectional study found that visual SL improves with age during childhood (between the ages of 5 and 12 years, Arciuli & Simpson, 2011). A similar pattern of age improvement was found for visual SL when both accuracy rates and hippocampal structure were compared in children, adolescents, and adults (Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017).

A recent study compared the developmental trajectory of visual and auditory SL across childhood (ages 5–12) and found modality-based differences: While visual SL improved with age, auditory SL did not (Raviv & Arnon, 2018). This pattern was taken as evidence for the modality-sensitive nature of SL, which is also supported by findings from the adult literature (Conway & Christiansen, 2005, 2006; Emberson et al., 2011; Raviv & Arnon, 2018; Siegelman & Frost, 2015). With regard to age, the conclusion was that auditory SL is age-invariant, whereas visual SL is not, a claim that can help reconcile the previously mixed findings in the literature where age effects were reported for visual SL (Arciuli & Simpson, 2011), but not auditory SL (Saffran et al., 1997). However, since the auditory task used linguistic stimuli (syllables), an alternative explanation is that the differential effect of age on performance was driven by the nature of the stimuli (linguistic vs. non-linguistic) rather than modality. Such a pattern would be consistent with the postulated role of SL in language acquisition, and the findings that unlike many other cognitive abilities, language learning does not improve with age (Hakuta, Bialystok, & Wiley, 2003; Hartshorne, Tennebaum & Pinker, 2018; Johnson & Newport, 3 1989). Linguistic auditory SL may similarly not improve with age. In other words, it is possible that auditory SL does improve with age when non-linguistic auditory stimuli are used.

In this study, we test this prediction by conducting a large-scale cross-sectional study of the developmental trajectory of visual and non-linguistic auditory SL across childhood (ages five to thirteen, $N = 232$). Infants, children, and adults are capable of learning non-linguistic auditory sequences and learn them as well as syllable sequences (Saffran, 2002; Saffran, Johnson, Aslin, & Newport, 1999). However, no study to date has asked whether performance improves with age for such non-linguistic auditory stimuli. We examine performance on a non-linguistic auditory task (using familiar sounds) and a visual one (using object drawings) across childhood. If the effect of age on SL is dependent on modality, reflecting constraints of the auditory system, we expect to see an improvement with age for the visual task, but not the auditory one, similar to Raviv and Arnon (2018). If, in contrast, the effect of age is stimuli-sensitive and age invariance is unique to linguistic elements, performance should improve with age for both the visual and non-linguistic auditory tasks. Our tasks are closely modeled on the ones used in Raviv and Arnon (2018), where participants had to detect recurring triplets in a continuous temporal stream. The task properties were similar to those of Raviv and Arnon (2018) in terms of their statistical properties, and nature and number of exposure and test trials. They differed in that our auditory task used familiar non-linguistic sounds (bird tweeting, door opening, and bell ringing etc.) instead of syllables: The sounds were familiar, but, unlike syllables, were not expected to co-occur together in any particular way. We tested children in the same age range as in the previous study to better enable the comparison between them.

## 2. Method

### 2.1. Participants

One hundred and eighteen children completed the auditory task (age range: 5–12 years, mean age: 8;7 years, 67 boys and 51 girls) and 111 different children completed the visual task (age range: 5–12 years, mean age: 8;7 years, 63 boys and 48 girls). All children were recruited as part of their visit to the Living Lab at the Bloomfield Science Museum in Jerusalem, and they received a small educational reward in return for their participation. Parental consent was obtained for all children. All the children were native Hebrew speakers, and none had known language or learning disabilities, as reported by their parents.

### 2.2. Materials

Each child completed either the non-linguistic auditory task or the visual task in a between-subject design. In both tasks, participants were exposed to a continuous exposure stream containing five recurring triplets. The transitional probabilities between syllables were 1 within triplets and 0.25 between triplets. Following exposure, learning was assessed using 25 two-alternative-forced-choice trials (2AFC).

### *2.2.1. Non-linguistic auditory task*

The auditory stimuli consisted of five unique triplets of familiar sounds built from an inventory of 15 different recognizable non-linguistic sounds such as door opening and bell ringing (see Appendix A for the complete list of sounds). The triplets were generated anew for each participant, so that each participant heard a different set of triplets. The duration of each sound was 500 ms, with a 100 ms break between sounds (duration of 1,800 ms for each triplet). This differs from the stimuli duration in Raviv and Arnon (2018) where each syllable was presented for 250 ms. This change allowed us to equate stimulus duration between the auditory and visual tasks, and to better reflect the longer duration of familiar sounds compared to syllables (we discuss the possible effect of this in the general discussion). The five triplets were concatenated together in a semi-randomized order to create the familiarization stream (with the constraint that no triplet will be heard twice in a row). The exposure phase lasted 3:30 min, with each triplet repeated 24 times. Importantly, there were no visual, auditory, or temporal cues to indicate triplet boundaries.

### *2.2.2. Visual task*

The visual stimuli consisted of five unique triplets of drawings from a set of 15 black-and-white drawings of familiar objects (house, book, plane; see Appendix A for all items). The drawings were all taken from a set of normed drawings (Alario & Ferrand, 1999). All pictures had high naming agreement, similar length (in Hebrew), high frequency, and early age of acquisition in Hebrew (based on Maital, Dromi, Sagi, & Bornstein, 2000). The triplets were created anew for each participant so that each saw a different set. The duration of each triplet was 1,800 ms (like the auditory task): Each drawing appeared on the screen for 500 ms, with a 100 ms break between drawings. The triplets were concatenated together in a semi-randomized order (with the constraint that no triplet will appear twice in a row) to create the visual familiarization stream. The exposure phase lasted 3:30 min, with each triplet repeated 24 times. Again, there were no visual, auditory, or temporal cues to indicate triplet boundaries.

### *2.2.3. The test-phase*

The test phase was identical for the two tasks. It included 25 2AFC trials where participants had to choose between two triplets (separated by 500 ms). On each trial, participants heard a real triplet (that appeared in the exposure stream) and a foil triplet, either preceding or following it. Foil triplets were constructed by taking the first sound/drawing from one triplet, followed by the second sound/drawing from another triplet, and the third sound/drawing from a third triplet. Each element in the foil triplets appeared in a similar position in real triplets, but with different surrounding sounds/drawings. This created a difference in TPs between real triplets and foils: While the TPs between every two adjacent elements within a triplet were 1, the TPs between every two elements in a foil were 0. If participants were attending to the distributional properties of the exposure stream, they should be able to distinguish between real triplets and foils.

### 2.3. Procedure

Children were tested individually while seated at a computer wearing noise-cancelling headphone. In the non-linguistic auditory task, children were told to listen to an alien song based on sounds they heard on Earth. Following exposure, children were asked to help the aliens reconstruct the song by telling them which of the two short sound sequences (triplets) appeared in the same order in the alien song they just heard. In the visual task, children were told that aliens were bringing souvenirs from Earth back to their spaceship. Following exposure, children were asked to help the aliens and say which of the two triplets they saw were taken into the spaceship in the same order before. For both tasks, the test trials were presented to children in random order (with the constraint that the same real/foil triplet did not appear in two consecutive trials). The order of real triplets and foils on each trial was counter-balanced so the real triplets appeared first in half of the trials and the foil appeared first in the other half. After hearing both possibilities, children were asked to press either "1" if they thought the correct triplet was the first or "2" if it was the second. In case children felt they did not know the answer, they were encouraged to guess which seemed more familiar. At the end of the task, the experimenter thanked the child for helping the aliens and let them to choose a small reward.

## 3. Results

Three children were removed from the analyses because their performance was significantly below chance using the binomial distribution (accuracy under 32% correct).[1] Children showed learning in both tasks with performance significantly above chance (non-linguistic auditory task: mean accuracy 59%, $t(118) = 8.36$, $SD = 0.12$, $p < .0001$; visual task: mean accuracy 64%, $t(111) = 8.96$, $SD = 0.16$, $p < .0001$). Performance was better in the visual task compared to the auditory task ($t(228) = 2.45$, $p < .01$), consistent with previous findings of a visual advantage in children of the same ages (Raviv & Arnon, 2018). Table 1 shows accuracy across development (in 5-year-and-a-half age bins of similar size, following Raviv & Arnon, 2018). We use these bins only for presentational clarity: In the analyses, age was entered in half years, without any additional division into bins. As can be seen, children showed learning as a group across development and seem to improve with age in both modalities, though the increase seems more linear in the visual domain (Table 1). Fig. 1 shows children's mean performance on both tasks as a function of age.

To test for significance, we used mixed-effect logistic regression models. Our dependent binominal variable was success in a single test trial. The model included fixed effects for age (in half years, centered), modality (visual vs. auditory. contrast coding),

Table 1
Visual and auditory SL accuracy by age bins

| | Non-linguistic Auditory | | | Visual | | |
|---|---|---|---|---|---|---|
| | *N* | Mean | *t*-score | *N* | Mean | *t*-score |
| Age group 5 to 6.5 | 24 | 0.58 | 3.49** | 22 | 0.54 | 1.25 |
| Age group 6.5 to 8 | 23 | 0.54 | 1.59 | 24 | 0.61 | 4.67*** |
| Age group 8 to 9.5 | 26 | 0.61 | 4.68*** | 24 | 0.64 | 4.33*** |
| Age group 9.5 to 11 | 25 | 0.57 | 3.54** | 24 | 0.68 | 5.18*** |
| Age group 11 to 12.5 | 24 | 0.64 | 6.02*** | 20 | 0.71 | 6.33*** |
| All children | 122 | 0.59 | 8.36*** | 114 | 0.64 | 9.3*** |

**$p < .01$; ***$p < .001$

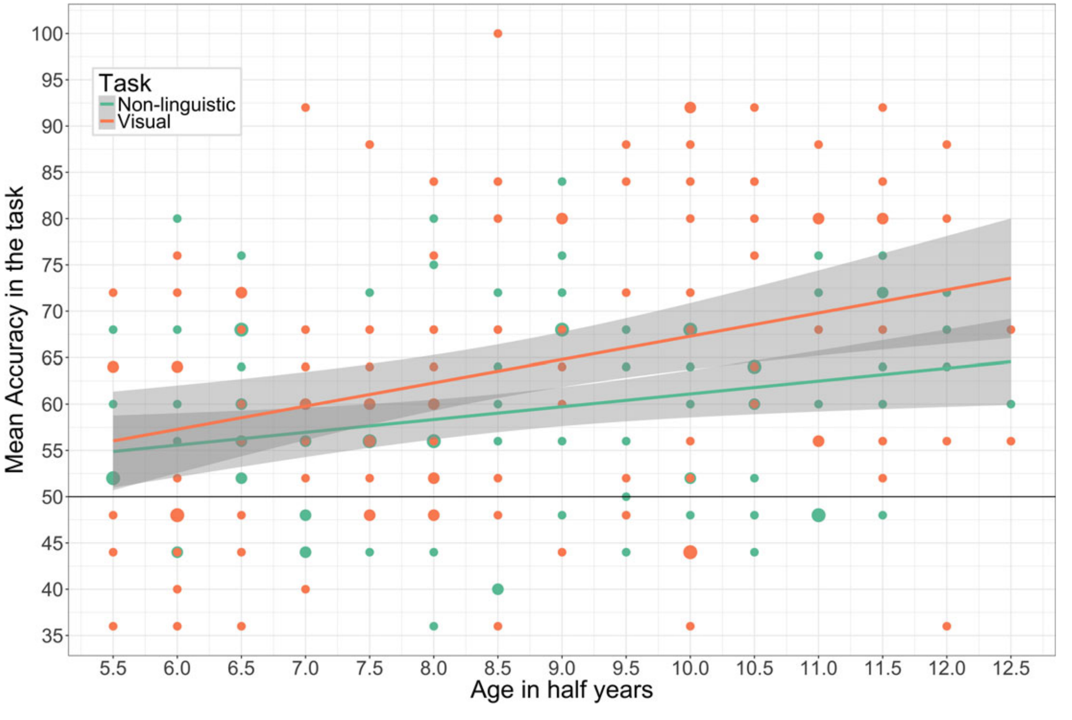


Fig. 1. Accuracy in both tasks by age (in half years). Each dot represents an accuracy score (ranging from 0% to 100%) shown by one participant in the relevant age range. The size of the dots reflects the number of participants who had that score. The two colorful plotted lines represent the linear regression line for each modality, with the standard confidence interval in black. The black line represents the 50% chance level.

and the interaction between them to see if age affects performance on both visual and non-linguistic auditory SL. We used age in half-years rather than age in months to better enable the comparison of our results to those of Raviv and Arnon (2018)—where age was used in half-years—and to avoid problems with model convergence (we are interested in the interaction between age and modality and do not have the same distribution of participants per month in the two tasks). The model also had trial number (centered), order of appearance in test (whether the real triplet or the foil appeared first, deviation coded), and gender (males vs. females, contrast coding) as additional fixed effects. We also included the interaction between order of appearance and modality, since previous work showed that order at testing affected accuracy for auditory, but not visual (Raviv & Arnon, 2018). The model also had a random intercept for subjects (because items were generated anew for each participant, we did not include a random effect structure for items).

Age had a significant positive effect on performance (Table 2): Children's accuracy improved with age, with older children showing higher accuracy ($\beta = 0.06$, $SE = 0.02$, $p = .032$). The interaction between modality and age was not significant ($\beta = 0.05$, $SE = 0.04$, $p = .16$, chi-square = 189, $p > .1$ in model comparison), indicating that auditory SL does improve with age when the stimuli are not linguistic. After removing the interaction term, the effect of age was highly significant ($\beta = 0.08$, $SE = 0.02$, $p < .001$, model comparison: chi-square = 18.03, $p < .001$). The effect of modality was significant, with children showing higher accuracy in the visual domain ($\beta = 0.20$, $SE = 0.07$, $p = .006$), as in Raviv and Arnon (2018). The effect of trial number was not significant ($\beta = -0.006$, $SE = 0.003$, $p = 0.09$), confirming that no learning (or unlearning) was happening during the test phase itself. The effect of gender was not significant ($\beta = -0.02$, $SE = 0.03$, $p = 0.49$). The effect of order of appearance was significant, with higher accuracy when the real triplet appeared first ($\beta = -0.008$, $SE = 0.03$, $p = .026$). The interaction between order of appearance and modality was marginal ($\beta = -0.09$, $SE = 0.05$, $p = .08$): The effect of order of appearance was not significantly stronger in the auditory modality.

We ran two additional models on each task separately to ensure that the effect of age is found in both modalities (Tables 3 and 4). We had the same fixed and random effect structure as in the previous model, excluding modality (Tables 2 and 3). Accuracy increased with age in both modalities (auditory: $\beta = 0.06$, $SE = 0.02$, $p = .005$; visual: $\beta = 0.11$, $SE = 0.03$, $p = .0007$). In line with the marginal interaction found in the joint model, the effect of order of appearance differed in the two modalities: Order of appearance affected accuracy in the auditory modality ($\beta = 0.08$, $SE = 0.03$, $p = .03$), but not

Table 2
Mixed-effect regression model for both tasks (ages 5–12) with significant effects in bold

| | Estimate | *SE* | *z*-value | *p*-value |
|---|---|---|---|---|
| **(Intercept)** | 0.373 | 0.054 | 6.866 | <.001 |
| **Age** | 0.06 | 0.028 | 2.134 | <.05 |
| **Modality (Visual)** | 0.209 | 0.078 | 2.677 | <.01 |
| Gender (male) | −0.02 | 0.039 | −0.655 | >.5 |
| Trial number | −0.006 | 0.003 | −1.676 | >.1 |
| **Order of appearance (first)** | 0.008 | 0.003 | 2.226 | <.05 |
| Age: modality (Visual) | 0.05 | 0.040 | 1.380 | >.15 |
| Modality (Visual): order of appearance | −0.09 | 0.055 | −1.729 | >.08 |

Table 3
Mixed-effect regression model for the auditory task (ages 5–12) with significant effects in bold

| | Estimate | *SE* | *z*-value | *p*-value |
|---|---|---|---|---|
| **(Intercept)** | 0.357 | 0.043 | 8.289 | <.001 |
| **Age** | 0.06 | 0.028 | 2.753 | <.01 |
| Gender (male) | −0.07 | 0.043 | −1.835 | =0.06 |
| Trial number | −0.004 | 0.005 | −0.854 | >.3 |
| **Order of appearance (first)** | 0.08 | 0.03 | 2.159 | <.05 |

Table 4
Mixed-effect regression model for the visual task (ages 5–12), significant effects in bold

| | Estimate | *SE* | *z*-value | *p*-value |
|---|---|---|---|---|
| **(Intercept)** | 0.611 | 0.069 | 8.781 | <.001 |
| **Age** | 0.11 | 0.034 | 3.390 | <.001 |
| Gender (male) | 0.03 | 0.069 | 0.554 | >0.5 |
| Trial number | −0.008 | 0.005 | −1.547 | >.1 |
| Order of appearance (first) | −0.01 | 0.041 | −0.317 | >.7 |

in the visual one ($\beta = -0.01$, $SE = 0.04$, $p = .75$), replicating previous findings (Raviv & Arnon, 2018).

### *3.1. Looking at the effect of age by modality and stimuli-type: A cross study comparison*

Our results show that both visual and auditory SL improve with age, when familiar sounds are used instead of syllables. This contrasts with previous findings (Raviv & Arnon, 2018) and suggests that the effect of age differs for linguistic and non-linguistic stimuli. To explore this claim more directly, we combined our data with that collected by Raviv and Arnon (2018) and compared the effect of age on the two auditory tasks (one linguistic and one non-linguistic). The two tasks had very similar exposure and testing properties. Each task contained five triplets, repeated 24 times, resulting in the same input statistics. The test phase was also similar, consisting of 25 2AFC trials. The only difference between the tasks was in the duration of the individual stimuli, with syllables having a shorter stimulus duration than the non-linguistic sounds (250 ms vs. 500 ms, we return to this difference in the discussion). We included all children between the ages of 6;6 and 12;0 in our analyses. We excluded children under 6;5 (for which we had data in both tasks) because they did not show learning in the linguistic auditory task (performance was at chance level). Including them leads to an effect of age for the linguistic task that only reflects the move from chance to learning (see Raviv & Arnon, 2018 for a more detailed explanation). This led to a total of 232 children: 119 children in the non-linguistic auditory task (mean age 8;7 years) and 113 in the linguistic one (mean age 8;5 years).

We ran a mixed effect regression model on the combined data with the same random and fixed effect structure used to compare the two modalities (see Table 5 and Fig. 2). We added stimuli-type (linguistic vs. non-linguistic) and the interaction between stimuli-type and age (in half-years) to test the prediction that age has a different effect on performance for linguistic and non-linguistic stimuli. In line with this prediction, the effect of age was not significant ($\beta = 0.003$, $SE = 0.03$, $p = .92$), but its interaction with stimuli-type was ($\beta = 0.08$, $SE = 0.04$, $p = .042$): Age increased accuracy only in the non-linguistic auditory task[2] . A simple-slope analysis confirmed this by showing that the effect of age was significant for the non-linguistic auditory task ($p = .002$), but not for the linguistic one ($p = .88$). To ensure that the interaction does not stem merely from combining two separate samples, we ran another model with the same effect structure on the combined data from our visual task and the one used in Raviv and Arnon (2018). The two visual tasks were identical in all exposure and testing properties and differed only in that one used black-and-white drawings and the other cartoon aliens. In contrast with the results for the auditory tasks, the effect of age was significant ($\beta = 0.11$, $SE = 0.03$, $p < .001$) while the interaction with stimuli-type (aliens vs. drawings) was not ($\beta = 0.062$, $SE = 0.04$, $p = .177$, see Table 6).

## 4. Discussion

In this study, we investigated the effect of modality on the developmental trajectory of SL during childhood. In particular, we wanted to test the hypothesis that age affects visual and auditory SL similarly, when non-linguistic auditory stimuli are used. To do so, we examined the developmental trajectory of visual SL and non-linguistic auditory SL (using sounds instead of syllables) in a large sample of children between the ages of 5 and 12. In contrast with previous findings (Raviv & Arnon, 2018), we found no difference in the effect of age on performance in the two modalities: Both visual and non-linguistic SL showed improvement during childhood (between the ages of 5 and 12), and the effect of age was not stronger in one modality. These results suggest that the linguistic nature of the stimuli (syllables), rather than its auditory modality, was responsible for

Table 5
Mixed-effect regression model for the two auditory tasks (linguistic and non-linguistic), ages 5–12, $N = 132$, with significant effects in bold

| | Estimate | *SE* | *z*-value | *p*-value |
|---|---|---|---|---|
| **(Intercept)** | 0.285 | 0.054 | 5.272 | <.001 |
| Age | 0.003 | 0.031 | 0.100 | >.9 |
| Stimuli-type (non-linguistic) | 0.020 | 0.007 | 0.270 | >.7 |
| Gender (male) | −0.063 | 0.034 | −1.860 | =.07 |
| Trial number | −0.006 | 0.004 | −1.513 | >.12 |
| **Order of appearance (first)** | 0.12 | 0.031 | 4.002 | <.001 |
| **Age: Stimuli-type (sounds)** | 0.124 | 0.031 | 2.208 | <.05 |

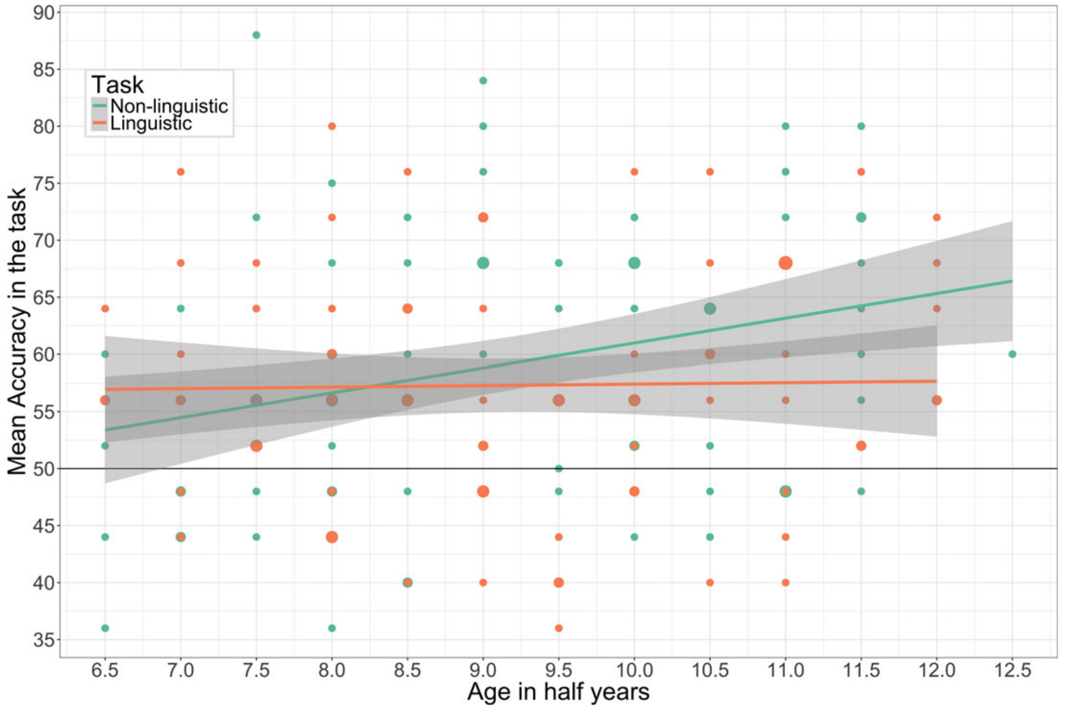


Fig. 2. Accuracy in the two auditory tasks by age (in half years). Each dot represents an accuracy score (ranging from 0% to 100%) shown by one participant in the relevant age range. The size of the dots reflects the number of participants who had that score. The two colorful plotted lines represent the linear regression line for each task, with the standard confidence interval in gray. The gray line represents the 50% chance level.

Table 6
Mixed-effect regression model for the two visual tasks (using aliens or drawings), ages 5–12, $N$ = 128, with significant effects in bold

| | Estimate | *SE* | *z*-value | *p*-value |
|---|---|---|---|---|
| **(Intercept)** | 0.595 | 0.065 | 9.091 | <.001 |
| **Age** | 0.11 | 0.03 | 3.578 | <.001 |
| Stimuli-type (aliens) | −0.072 | 0.0091 | −0.794 | >.4 |
| Gender (male) | −0.015 | 0.045 | −0.332 | >.7 |
| Trial number | −0.005 | 0.003 | −1.361 | >.16 |
| Order of appearance (first) | 0.015 | 0.028 | 0.529 | >.5 |
| Age: stimuli-type (aliens) | 0.015 | 0.062 | 1.438 | >.17 |

the lack of improvement with age in previous work. To explore this claim more directly, we combined our data with that collected by Raviv and Arnon (2018) and compared the effect of age on the two auditory tasks (one linguistic and one non-linguistic). This comparison revealed a significant interaction between age and stimuli type in the auditory domain: Accuracy increased with age only for the non-linguistic auditory task. To ensure the interaction was not driven by combining two separate samples, we also compared the effect of age on the two visual tasks (one using drawings and the other alien cartoons) and found no interaction: Age improved performance similarly on both visual tasks. Taken together, our findings support the hypothesis that the effect of age on performance is modulated by stimuli-properties, not modality. They further suggest that age invariance is limited to linguistic stimuli: SL improved with age across modality when non-linguistic stimuli were used.

Before reaching this conclusion, it is important to consider additional differences between the stimuli used in each study. In particular, the two auditory tasks differed in stimulus duration: Each syllable was presented for 250 ms, while each familiar sound (and the visual images) were presented for 500 ms. The difference in duration raises a potential challenge to our interpretation: The lack of age effects for the linguistic stimuli could reflect their shorter duration, rather than their linguistic nature. That is, the positive effect of age could result from an improvement in other skills, such as strategic processing, which develop with age and may be less available at the shorter presentation rate. Several factors undermine this alternative explanation. First, such an explanation predicts that SL will not improve with age at shorter durations. This contrasts with recent developmental data showing that children's visual SL improves similarly with age for shorter and longer presentation rates (200 ms, 400 ms, and 800 ms; Arciuli & Simpson, 2011). It could still be that such an interaction exists only in the auditory domain, but this would be surprising given that changes in duration impact visual SL more than auditory SL, at least in adults (Conway & Christiansen, 2009; Emberson et al., 2011). Second, both presentation rates were relatively long (even 250 ms is not a rapid presentation rate for a single syllable), and the difference between them is not one that is implicated (to our knowledge) in the activation—or lack thereof—of relevant cognitive processes like memory or attention. Finally, findings from adults suggest that some explicit strategies are available even at short stimulus durations (250 ms; Bertels et al., 2015), and more important, that their impact on incidental learning may be limited, even for longer durations. Longer presentation rates led to improved performance only for intentional learning (when participants were told in advance that shapes appear in recurring triplets), and not for incidental learning like the one we assessed (Bertals, Destrebecqz, & Franco, 2015). While it is unlikely that the difference between the two auditory tasks was driven solely by duration, further work is needed to disentangle the impact of stimulus duration and stimulus type on the developmental trajectory of SL (we are currently examining the effect of age on non-linguistic stimuli presented for a shorter duration).

There are several possible explanations for the differential effect of age on linguistic and non-linguistic auditory stimuli. One possibility is that learning linguistic relations in general is age-invariant; however, this seems implausible given that declarative memory—which is implicated in various aspects of word learning—improves during childhood (e.g., Ofen et al., 2007). Alternatively, the lack of age effects may be limited to the kind of syllable statistics tested. Learning about syllable co-occurrence patterns—the relation we tested and that is tested in many auditory SL tasks—is necessary for word segmentation, one of the first challenges of language acquisition. Learning these particular relations may not change much after the first year of life, when phonetic categories are established and word segmentation is performed (e.g., Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). Linguistic auditory tasks are sensitive to prior knowledge in a way that other SL tasks are not (Perruchet, Poulin-Charronnat, Tillmann, & Peereman, 2014; Poulin-Charronnat, Perruchet, Tillmann, & Peereman, 2017; Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018), but the effect of that knowledge may not change much after it has been established. Since knowledge about syllable co-occurrences is already present in early childhood (Storkel, 2001), our youngest age group (six-year-olds) may have been too “old” to show age effects. That is, age may still affect linguistic SL, but only at younger ages. Testing this prediction requires evidence for the entire age range from birth to later childhood, which is extremely hard to come by. Research on SL in children has tended to adopt the same methods used with adults, where learning is assessed using explicit measures (e.g., Kidd & Arciuli, 2016; Saffran et al., 1997). These measures are not suitable for use with younger children (and are not very reliable with older children as well, Arnon, under review). There is currently no existing SL task that can be used across the entire age range.

The current findings also have implications for our understanding of the domain-generality versus domain-specificity of SL. The fact that SL is found across modalities has been used to argue for its domain-generality (Saffran & Thiessen, 2007; Thiessen & Erickson, 2015). However, there is growing evidence that SL shows modality-specific constraints and may function differently across modalities (Arciuli, 2017; Conway & Christiansen, 2005, 2006; Emberson et al., 2011; Frost, Armstrong, Siegelman, & Christiansen, 2015; Siegelman & Frost, 2015). Our findings highlight *similarities* between visual and auditory SL and point to the importance of taking stimuli properties into account when investigating the effect of modality. Previously reported differences (and similarities) may have been driven not only by modality, but also by the specific stimuli used. For instance, stimuli choice may modulate overall differences in accuracy across modalities. Adults show better learning in the auditory domain compared to the visual one (Conway & Christiansen, 2005; Emberson et al., 2011; Siegelman & Frost, 2015), a finding taken to reflect the greater sensitivity of audition to temporal regularities. Children, in contrast, show better learning in the visual domain (Raviv & Arnon, 2018; the current paper). Rather than reflecting a developmental stage, this discrepancy may stem from the different stimuli used in child and adult studies. The auditory advantage in adults is found when comparing familiar syllables (auditory) to unidentifiable shapes (visual). The visual advantage in children is found when comparing syllables (auditory) to cartoon aliens or familiar drawings: In these cases, the visual stimuli is probably easier to encode and/or remember (it has more semantic content). The existence (and direction) of a modality-based advantage may depend on the relative familiarity and salience of the particular visual and auditory stimuli used. More generally, claims about modality-based differences in SL need to be validated for a range of linguistic and non-linguistic stimuli. The effect of stimuli choice on performance highlights the impact of prior knowledge and expectations on SL (Siegelman et al., 2018) and is consistent with a view of SL as a multi-component mechanism (Arciuli, 2017). The computations themselves may be domain-general, and work similarly across modality and stimuli, but nevertheless result in varying outcomes (and developmental trajectories) because of the difference in the existence and strength of prior knowledge. Under this view, the "uniqueness" of linguistic stimuli is not related to its auditory nature but reflects the stronger and more entrenched expectations we have about the temporal regularities of speech. Similar effects could be found in the processing of visual stimuli for which we have similarly strong expectations (like letter patterns).

In sum, this study compared the developmental trajectory of visual and non-linguistic auditory SL. The results show that both improved similarly with age during childhood, in contrast to linguistic auditory SL, which was age-invariant for the same age range (Raviv & Arnon, 2018). While additional work is needed to understand the effect of stimulus duration on SL during development, these findings highlight the stimuli-sensitive nature of SL and suggest that previously found differences were not driven solely by modality. As such, they are consistent with the view that SL is not a unitary, stable capacity, but rather one that is sensitive to both stimulus and modality features.

## Acknowledgments

We thank Noam Siegelman, Louisa Bogaerts, and Ram Frost for helpful comments and discussions, and Limor Raviv for sharing her data. We thank Zohar Aizenbud for programming the experiments. We also thank the research assistants at the Living Lab in the Bloomfield Science Museum (Tamar Johnson, Ori Lavi-Rotbein, Niva Goldberg); the museum staff; and the children and parents who participated in the studies. The research was funded by an Israeli Science Foundation grant to the second author (grant number 584/16).

## Notes

1. These children were removed because they are showing a clear dispreference for the familiar items, suggesting they did not understand the forced-choice task. The results do not change is they are included.
2. Based on Emberson et al. (2011) we would have expected performance on the linguistic task to be higher overall than the non-linguistic task because of its' shorter duration, but this was not the case. It is hard to directly compare the two studies since the intervals between stimuli in Emberson et al. (2011) were much longer than in the current study: Because they used interleaved visual and auditory streams, their long presentation rate, where auditory SL deteriorated, had a gap of up to 2,250 m between individual elements (see p. 1029 in that paper). It may be that auditory SL only deteriorates at very slow presentation rates, slower than the one we used

## References

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 531–552. https://doi.org/10.3758/BF03200732.

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160058. https://doi.org/10.1098/rstb.2016.0058.

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x.

Arciuli, J., & von Koss Torkildsen, J. (2012). Advancing our understanding of the link between statistical learning and language acquisition: The need for longitudinal data. *Frontiers in Psychology*, *3*, 1–9. https://doi.org/10.3389/fpsyg.2012.00324.

Arnon, I. (under review). Do current statistical learning tasks capture stable individual differences in children? An investigation of reliability across modalities.

Bertels, J., Boursain, E., Destrebecqz, A., & Gaillard, V. (2015). Visual statistical learning in children and young adults: How implicit? *Frontiers in Psychology*, *5*(1541), 1–11. https://doi.org/10.3389/fpsyg.2014.01541.

Bertels, J., Destrebecqz, A., & Franco, A. (2015). Interacting effects of instructions and presentation rate on visual statistical learning, *Frontiers in Psychology*, *6*: 1806, https://doi.org/10.3389/fpsyg.2015.01806

Christiansen, M.H. (in press). Implicit-statistical learning: A tale of two literatures. *Topics in Cognitive Science*

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*(3), 356–371. https://doi.org/10.1016/j.cognition.2009.10.009.

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. https://doi.org/10.1037/0278-7393.31.1.24.

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. https://doi.org/10.1111/j.1467-9280.2006.01801.x.

Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, *21*, 561–580.

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *The Quarterly Journal of Experimental Psychology*, *64*(5), 1021–1040. https://doi.org/10.1080/17470218.2010.538972.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010.

Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, *57*(1), 56–74. https://doi.org/10.1016/j.cogpsych.2007.12.001.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. https://doi.org/10.1016/S0010-0277(99)00003-7.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31–38. https://doi.org/10.1111/1467-9280.01415.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. https://doi.org/10.1037/a0025405.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193. https://doi.org/10.1111/cdev.12461.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608. https://doi.org/10.1126/science.1736364.

Lukács, Á., & Kemény, F. (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*, *39*(2), 383–404. https://doi.org/10.1111/cogs.12143.

Maital, S. L., Dromi, E., Sagi, A., & Bornstein, M. H. (2000). The Hebrew communicative development inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, *27*(1), 43–67.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111. https://doi.org/10.1016/S0010-0277(01)00157-3.

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*(3), 199–221. https://doi.org/10.1006/jecp.1998.2442.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302–331. https://doi.org/10.1111/j.1467-9922.2010.00626.x.

Ofen, N., Kao, Y.-C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, *10*, 1198–1205. https://doi.org/10.1038/nn1950.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006.

Perruchet, P., Poulin-Charronnat, B., Tillmann, B., & Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta Psychologica*, *149*, 1–8. https://doi.org/10.1016/j.actpsy.2014.01.015.

Poulin-Charronnat, B., Perruchet, P., Tillmann, B., & Peereman, R. (2017). Familiar units prevail over statistical cues in word segmentation. *Psychological Research*, *81*(5), 990–1003. https://doi.org/10.1007/s00426-016-0793-y.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*, https://doi.org/10.1111/desc.12593.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford, UK: Oxford University Press.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172–196. https://doi.org/10.1006/jmla.2001.2839.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and Learning) out of the comer of your ear. *Psychological Science*, *8*(2), 101–105. https://doi.org/10.1111/j.1467-9280.1997.tb00690.x.

Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.), *Blackwell handbooks of developmental psychology. Blackwell handbook of language development* (pp. 68–86). Malden, MA: Blackwell. https://doi.org/10.1002/9780470757833.ch4

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, *29*(1), 37–51. https://doi.org/10.1162/jocn_a_01028.

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Statistical entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198–213. https://doi.org/10.1016/j.cognition.2018.04.011.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, *44*, 1321–1337. https://doi.org/10.1044/1092-4388(2001/103).

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*(1), 21. https://doi.org/10.1186/1471-2202-10-21.

Thiessen, E. D., & Erickson, L. C. (2015). Perceptual development and statistical learning. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 396–414). New York: Wiley-Blackwell.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3.

## Appendix A: List of stimuli used for both tasks

1. Non-linguistic auditory task (familiar sounds): "bird tweet," "running water," "goat bleat," "opening door," "dog bark," "bouncing ball," "trumpet," "cat meow," "duck quack," "frog quack," "cuckoo clock," "bell," "chord," "whistle," and "cow moo."
2. Visual task: black-and-white drawings (taken from Alario & Ferrand, 1999) of house, book, cake, dog, cat, airplane, shoe, fish, ball, banana, fork, flower, bottle, chair, and butterfly.

Example images: