Cognition 206 (2021) 104492

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Short Communication

Visual statistical learning is facilitated in Zipfian distributions

Ori Lavi-Rotbain [a,*], Inbal Arnon [b]

[a] The Edmond and Lilly Safra Center for Brain Sciences, Hebrew University, Israel
[b] Department of Psychology, Hebrew University, Israel

ARTICLE INFO

Keywords:
Visual statistical learning
Zipfian distribution
Information theory
Predictability
Domain-general
Learning

ABSTRACT

Humans can extract co-occurrence regularities from their environment, and use them for learning. This statistical learning ability (SL) has been studied extensively as a way to explain how we learn the structure of our environment. These investigations have illustrated the impact of various distributional properties on learning. However, almost all SL studies present the regularities to be learned in uniform frequency distributions where each unit (e.g., image triplet) appears the same number of times: While the regularities themselves are informative, the appearance of the units cannot be predicted. In contrast, real-world learning environments, including the words children hear and the objects they see, are not uniform. Recent research shows that word segmentation is facilitated in a skewed (Zipfian) distribution. Here, we examine the domain-generality of the effect and ask if visual SL is also facilitated in a Zipfian distribution. We use an existing database to show that object combinations have a skewed distribution in children's environment. We then show that children and adults showed better learning in a Zipfian distribution compared to a uniform one, overall, and for low-frequency triplets. These results illustrate the facilitative impact of skewed distributions on learning across modality and age; suggest that the use of uniform distributions may underestimate performance; and point to the possible learnability advantage of such distributions in the real-world.

1. Introduction

Much work over the past decades has examined humans' ability to extract regularities from their environment as a way to explain how we detect structure within the "buzzing, blooming confusion" (James, 1890) that the real world. Often called statistical learning, the ability to detect co-occurrence regularities and use them to learn higher order structure is found across modalities (Kirkham, Slemmer, & Johnson, 2002) is present from early on (Bulf, Johnson, & Valenza, 2011), and plays an important role in early development (see Saffran & Kirkham, 2018 for a review). Studies of SL illustrate learners' sensitivity to the structure of the environment, and provide a way to examine the factors in the environment that learners are sensitive to. One much studied example is learners' sensitivity to the transitional probabilities (TPs) between elements. In natural language, TPs are higher between syllables within the same word compared to across word boundaries. Accordingly, infants, children, and adults are capable of using TPs to detect word boundaries and segment a novel speech stream (e.g., Saffran, Aslin, & Newport, 1996). The effect of TPs is not limited to the auditory domain: TPs can also be used to segment a stream of recurring visual triplets (Kirkham et al., 2002). These studies illustrate how the distributional information present in the environment – in this case TPs - is utilized during learning and helps learners detect recurring units in the input.

Another, less studied, characteristic of our real-world environment is that some units tend to appear more often than others. This is most famously noted for words in language, whose frequency follows a Zipfian distribution showing a power law relation between a words' frequency and its' rank (Piantadosi, 2014; Zipf, 1949). Intuitively, this reflects the fact that language has few high frequency words, many low frequency words, and that the decrease in word frequency is not linear (e.g., the first word is twice as frequent as the second, and so on). This skewed word distribution is also found in the input children hear, making it a consistent feature of their linguistic learning environment (Hendrickson & Perfors, 2019; Lavi-Rotbain & Arnon, 2020b). Interestingly, the objects children see also have a right-skewed distribution, with few objects appearing very often and many appearing rarely (Clerkin, Hart, Rehg, Yu, & Smith, 2017). Such power law distributions are common across the physical world (e.g., Newman, 2005), suggesting that the environment whose structure children need to learn is often

* Corresponding author at: The lab of Prof. Inbal Arnon, Psychology Department, The Hebrew University, Mt. Scopus, Jerusalem 9190501, Israel.
E-mail address: oriedit.lavi@mail.huji.ac.il (O. Lavi-Rotbain).

https://doi.org/10.1016/j.cognition.2020.104492
Received 4 February 2020; Received in revised form 12 October 2020; Accepted 20 October 2020
Available online 3 November 2020
0010-0277/

skewed in a particular way.

Despite their recurrence in the real-world, relatively little work has examined the impact of skewed distributions on learning. Most SL studies use a uniform distribution of elements, where each unit to be learned appears the same number times. While the use of uniform distributions is useful because it allows us to control for many factors (i.e. frequency), such distributions differ from real-world environments and may be harder to learn form. Specifically, skewed distributions may confer a learnability advantage because they are more predictable than uniform distributions where each element is equally likely to appear. This increased predictability could assist learning by making it easier to predict upcoming elements (thereby freeing up processing resources), and by allowing learners to learn high frequency elements early on, and use them to facilitate learning of lower frequency elements. To take an example, high frequency words can be identified quickly, and used as an anchor to segment lower frequency words, as seen in infants' use of their own name to segment adjacent words (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). In line with this postulated learnability advantage, a handful of recent studies show that learning in several linguistic tasks benefits from exposure to skewed distributions. Word segmentation was improved for both children and adults after exposure to a Zipfian distribution compared to a uniform one (Kurumada, Meylan, & Frank, 2013; Lavi-Rotbain & Arnon, 2019b; Lavi-Rotbain & Arnon, 2020a). Similarly, learning of novel object-label associations via cross-situational learning (Hendrickson & Perfors, 2019), and through the integration of multimodal cues (Lavi-Rotbain & Arnon, 2019a) was also improved in a Zipfian distribution. On a more abstract level, grammatical category learning was preserved under a Zipfian distribution, despite the overall lower frequency of the words in each category (Schuler, Reeder, Newport, & Aslin, 2017).

The few studies that have examined learning from Zipfian distributions suggest they are beneficial for learning, but the extent and generality of this effect is still unclear. For starters, the number of studies documenting such an effect is very small. More importantly, the existing evidence is limited to the linguistic domain. Here, we examine the domain-generality of the effect by investigating the propensity of skewed distributions in the visual domain, and their impact on learning visual co-occurrence regularities. Specifically, we ask whether visual SL - detecting recurring object triplets in a continuous stream - will be facilitated in a Zipfian distribution. While the Zipfian nature of the linguistic input is well-documented, very little work has asked whether other aspects of children's environment are also skewed. Interestingly, a recent study suggests that skewed distributions are also found in the visual domain: Clerkin et al. (2017) analyzed the distribution of single objects in infant's visual field and found that they follow a Zipfian distribution, with consequences for learning: The frequency rank of an object was predictive of learning its' label (the names for more frequent objects were learned earlier).

These findings suggest that children's visual input is skewed, and that this natural skew impacts learning trajectories. This investigation, however, is limited to the distribution of single objects, and does not tell us whether the co-occurrence of objects is also skewed. Humans' ability to detect and learn visual co-occurrence patterns has been studied extensively within the visual SL literature (e.g., Kirkham et al., 2002). However, all prior studies have used uniform distributions where each image triplet or pair appear equally often. Such distributions may not be representative of the structure of the visual environment and may negatively impact learning: If the co-occurrence of objects is also skewed, and if exposure to more predictable distributions facilitates learning across modality, we would expect visual SL to be improved in a Zipfian distribution. Such findings would extend our understanding of the real-world visual environment and suggest that the effect of more predictable distributions on learning is domain general. They would highlight an additional factor that impacts SL (distribution shape), and the importance of taking this into account when designing and interpreting SL studies. Looking at the impact of skewed distributions on visual SL is also relevant for understanding the domain-generality of statistical learning itself (e.g., Frost, Armstrong, Siegelman, & Christiansen, 2015). While SL is found across modalities (Saffran & Kirkham, 2018), there is evidence it is affected by stimuli type. In particular, SL of linguistic stimuli seems to differ from that of non-linguistic stimuli - visual or auditory - in several respects: Linguistic SL is more affected by prior knowledge (Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018); it does not change developmentally in the same way that visual and non-linguistic SL do (Raviv & Arnon, 2018; Shufaniya & Arnon, 2018); and individual performance on linguistic SL does not correlate with performance on non-linguistic tasks (though the non-linguistic tasks do correlate with each other, Arnon, 2020; Siegelman & Frost, 2015). That is, it is not immediately clear that an effect found in the linguistic domain will extend to the visual one.

In the current study we explore the propensity of skewed distribution in the visual domain and their impact on learning. We extend prior work by showing that the distribution of object pairs and triplets is also skewed, and that exposure to a skewed environment facilitates learning the regularities between visual elements. We start by looking at the visual input itself: we examine the distribution of object pairs and triplets in the learning environment using the same database used in Clerkin et al., 2017 (available at: https://nyu.databrary.org/volume/268). We find that both object pairs and triplets have a near-Zipfian distribution. We then use a classic visual SL paradigm to ask whether learning the co-occurrence patterns between objects is facilitated in a similarly skewed distribution in older children (mean age 10;9 years) and adults. The child sample serves as a replication in another population of learners, and an extension: Given that visual SL improves with age (Raviv & Arnon, 2018), we wanted to see if the effect of distribution type holds despite children's lower overall performance. We assess the impact of distribution type by comparing learning of six recurring visual triplets in a uniform distribution where each triplet appears 24 times; and in a Zipfian distribution, where triplets vary in frequency according to a power law distribution (with triplets appearing 92, 22, 12, 8, 6 or 4 times). In both distributions, participants were exposed to an unsegmented stream of black and white drawings, containing recurring triplets. Following exposure, participants were asked to differentiate between real triplets and foils. We expect performance in the Zipfian distribution to be better than in the uniform distribution, overall, and for the lower frequency triplets (despite them appearing less than in the uniform distribution).

2. Study 1: do objects combinations have a skewed distribution in the real world?

Here, we examine the distribution of two- and three-object combinations in the learning environment. We focus on object pairs and triplets since these are the most commonly used stimuli in the experimental visual SL literature (Frost, Armstrong, & Christiansen, 2019). We used the same dataset used in Clerkin et al. (2017, https://nyu.databrary.org/volume/268). In their study, Clerkin et al. used a head mounted camera to capture what infants see during mealtimes. 5775 of the collected images were tagged by adults via Amazon's MTurk platform for the visual objects appearing in them, generating a list of unique objects for each image (see the original paper for details). We used this dataset to create a list of all the object pairs and triplets that appeared within the same image (e.g., for an image with ten unique objects, there would be 45 possible object pairs and 120 possible object triplets). We then calculated how many times each of them appeared across the entire dataset. We used the same cleaning-up procedures described in the original paper (e.g., correcting spelling errors; changing plurals to singulars and removing adjectives), resulting in a dataset with the same number of individual objects as in the original paper. The dataset yielded 20,064 unique object pairs and 137,247 unique object triplets.

As was found for individual objects (Clerkin et al., 2017), object pairs and triplets (Fig. 1) also followed a right skewed distribution. Some

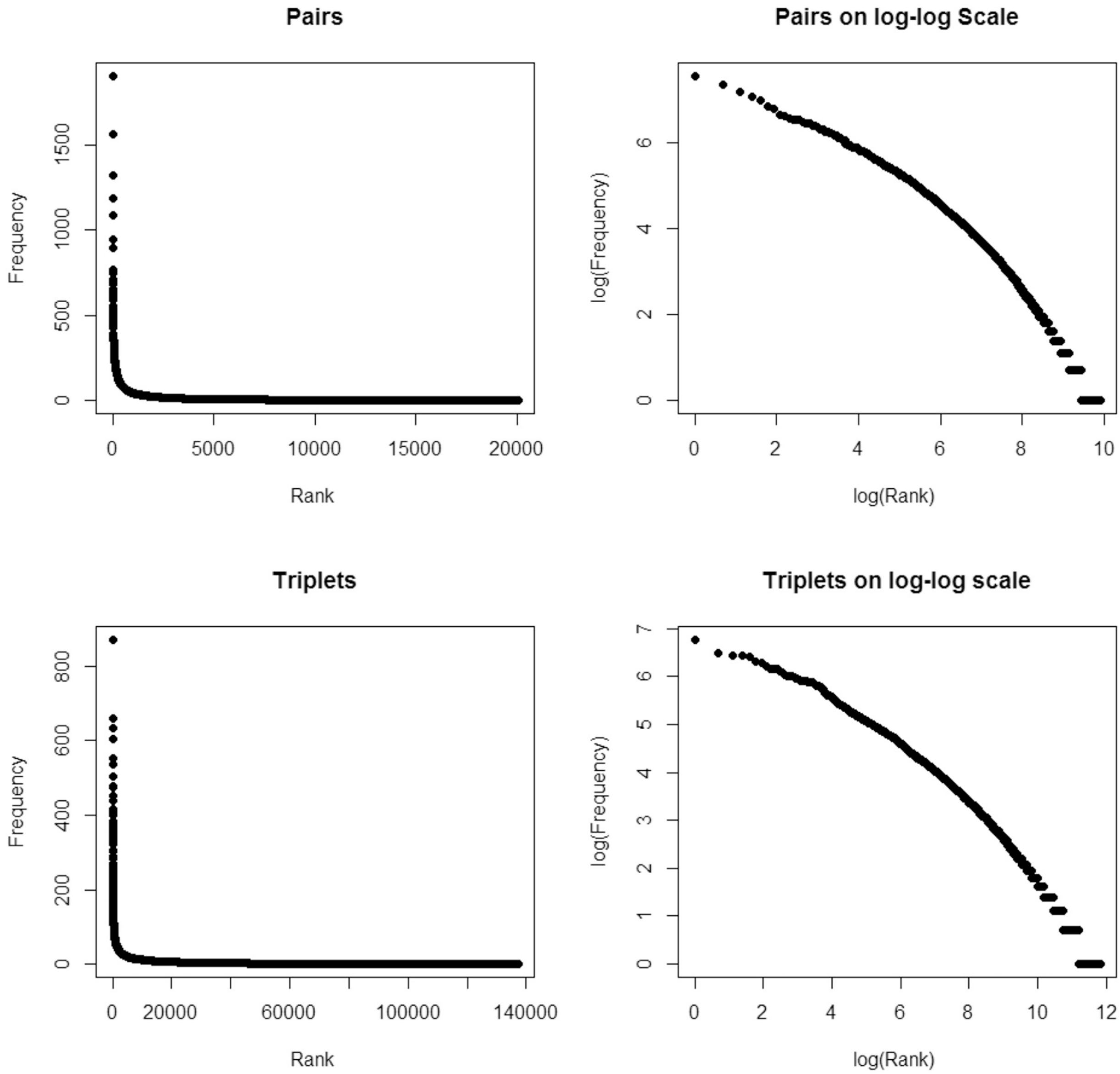


Fig. 1. The statistics of object pairs (top half) and triplets (bottom half) are presented on a frequency-rank scale (left) and on a log(frequency)-log(rank) scale (right). The distributions of pairs and of triplets are right skewed.

object combinations appeared frequently, while most appeared rarely. The distribution was near-Zipfian,as reflected in the high fit of a linear line to the log space (R^2(pairs) = 0.96; R^2(triplets) = 0.962; a fit comparable to the one reported for single objects in Clerkin et al. (2017)). That is, despite not being traditionally characterized as such, object co-occurrence patterns have a skewed distribution in children's input., suggesting that the relation between objects in children's visual environment is more predictable than previously thought. Next, we ask whether children can take advantage of this skew during learning.

3. Study 2: visual SL is facilitated in a Zipfian distribution

After seeing that objects combinations have a skewed frequency distribution in the learning environment, we test the prediction that visual SL will be improved after exposure to a Zipfian distribution compared to a uniform one. We look at children and adults in the same task to see if the effect of distribution type holds across age and modality. We focus on children in the same age range (9;0–12;0) where a facilitative effect was found in a linguistic SL task (Lavi-Rotbain & Arnon, 2019a).

4. Method

4.1. Participants

124 participants participated in the study: 64 undergraduate students (46 females, 18 males, M_{age} = 24;0) and 60 children (age range: 9;0–12;11 years, Mean age: 10;9 years; 24 boys, 36 girls). Children were visitors at the Bloomfield Science Museum in Jerusalem and were recruited as part of their visit to the Living Lab. Parental consent was obtained for all children. All participants were native Hebrew speakers without learning disabilities or attention deficits. Adult participants received 10 NIS or course credit in return for their participation and children received a small prize.

4.2. Materials

The task was closely modelled on the one used in Shufaniya and Arnon (2018). Participants were exposed to a familiarization stream containing a continuous stream of visual images. The stream contained six unique triplets created from a set of 18 black- and-white drawings of familiar objects (e.g., dog, shoe, plane; see Appendix A for all items). The 18 images were chosen from a normed dataset (Alario & Ferrand, 1999b) to have high naming-agreement; similar syllable length in Hebrew; high Hebrew frequency; and early Hebrew age-of-acquisition (based on Maital, Dromi, Sagi, & Bornstein, 2000). Triplets were created anew for each participant. During exposure, each image was presented for 500 ms, followed by a blank screen (100 ms, see Fig. 2). The total length of exposure for each triplet was 1800 ms.

In the uniform condition, each triplet appeared 24 times in a semi-randomized order, with the constraint that no triplet will appear twice in a row. In the Zipfian condition, the frequency of each triplet varied following a power law distribution (92, 22, 12, 8, 6 or 4 appearances, see Table 1). The exposure stream was composed of two blocks created for each participant by randomly sampling triplets according to their frequency, to ensure that the infrequent triplets appeared throughout the experiment. Importantly, participants were not aware of this division in any way: There was no break between the blocks and learners could not tell when one block ended and the other started. Because triplets were created anew for each participant, the frequent triplet differed across participants, reducing item-specific effects. The total number of appearances and the duration were identical in the two conditions ($N = 144$ repetitions, duration 4:30 min). In both conditions, the only cue for triplet boundaries were the higher TPs between images within triplets (TP = 1), compared to across triplets. This cue was present and similar in both distributions (0.2 in the uniform condition; 0.167 on average in the Zipfian condition).

4.2.1. *Procedure*

Participants completed the experiment on a computer while seated in a quiet room. They were randomly assigned to one of the two conditions, and were given identical instructions. Participants were told that aliens had visited Earth and taken several objects back with them. They were asked to pay attention to the order objects appeared in while listening to non-vocal music. Children completed the experiment while sitting next to an experimenter. After exposure, participants completed a segmentation task containing 36 two-alternative forced choice trials. On each trial, participants saw a triplet and a foil (in counter-balanced order) separated by 500 msec. Participants were asked to choose the triplet where the order of items was like what they saw during exposure. Six foils were created by mixing three images from three different triplets, while maintaining their original position (taking the first image from one triplet, the second from another, and the third from another). The TPs between images in the foils were always 0. Each real triplet was matched with each foil, resulting in 36 trials. The trials were randomized for each participant, with the constraint that the same word/foil did not appear in two consecutive trails.

5. Results

There were 64 participants in the uniform condition (34 adults and 30 children) and 60 in the Zipfian condition (30 adults and 30 children).

Table 1
Different experimental conditions.

Condition	Length [min]	Total exposure	No. repetitions per triplet	TPs (all 1 within words)
Uniform	4:30	144	24	Between: 0.2
Zipfian	4:30	144	Frequent: 92 Infrequent: 22, 12, 8, 6, 4	Between: 0.167 on average

Despite the random sampling, children in the uniform condition were slightly older than children in the Zipfian condition ($M_{age\text{-}uniform} = 11;1$, $M_{age\text{-}Zipfian} = 10;6$, $t(57.96) = 1.96$, $p = 0.055$). However, this difference goes against our prediction since older children show better visual SL (Raviv & Arnon, 2018) while we predict better performance in the Zipfian condition.

Children and adults showed learning in both conditions (were above chance, uniform: $M_{adults} = 83.33\%$, $t(33) = 12.81$, $p < 0.001$; $M_{children} = 70.18\%$, $t(29) = 8.06$, $p < 0.001$; Zipfian: $M_{adults} = 90.93\%$, $t(29) = 19.25$, $p < 0.001$, $M_{children} = 77.04\%$, $t(29) = 8.45$, $p < 0.001$). As predicted, both showed better learning in the Zipfian condition compared to the uniform one (see Fig. 3, see Table B.1 in Appendix B for accuracy by triplet frequency).

We used general linear regression models to examine the effect of distribution type on performance. Following Barr, Levy, Scheepers, and Tily (2013), the models had the maximal random effect structure justified by the data that would converge. Our dependent binominal variable was accuracy on a single trial of the segmentation test. Our fixed effects were: distribution type (uniform as baseline); age group (adult or child); triplet log frequency; trial number (centered); and order-of-appearance in the test (real-triplet-first vs. foil-first). We included the interaction between distribution type and age group to make sure that the effect of distribution type on accuracy was similar for both age groups. The model had random intercepts for participants, but not for items because triplets were generated anew for each participant (see Table 2). To examine the overall effect of experimental condition we used model comparisons.

As predicted, participants were more accurate in the Zipfian condition compared to the uniform one ($\beta = 0.99$, $SE = 0.33$, $p < 0.01$; chi-square(2) = 13.32, $p < 0.01$). Children were less accurate overall, ($\beta = -0.1$, $SE = 0.3$, $p < 0.001$), unsurprising since visual SL improves with age (Raviv & Arnon, 2018). Importantly, the interaction between condition and age group was not significant ($\beta = 0.31$, $SE = 0.44$, $p > 0.1$), meaning that learning was facilitated in the Zipfian distribution similarly for both children and adults. Accuracy was higher for more frequent triplets ($\beta = 0.13$, $SE = 0.05$, $p < 0.01$). Trial number affected performance, with better accuracy earlier on ($\beta = -0.02$, $SE = 0.004$, $p < 0.001$), as has been previously found (Raviv & Arnon, 2018). Order-of-appearance during test did not affect performance ($\beta = -0.09$, $SE = 0.08$, $p > 0.1$), as has been found before in the visual domain (Raviv & Arnon, 2018).

To ensure that the advantage in the Zipfian condition was not driven

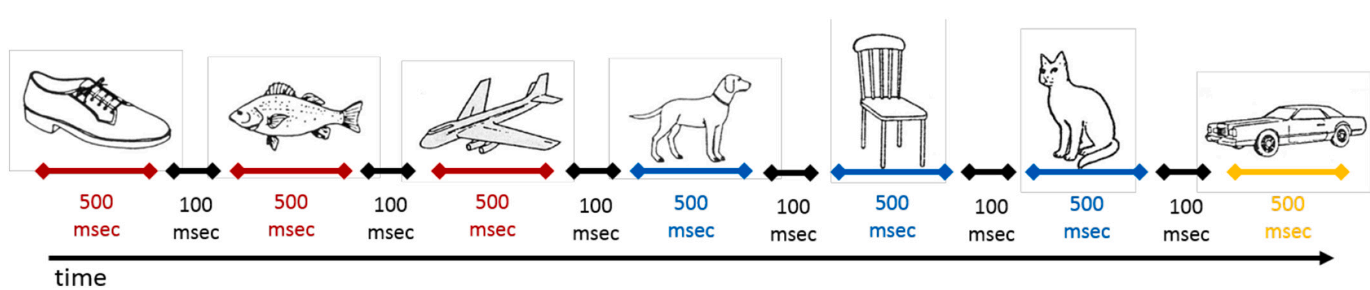


Fig. 2. Exposure illustration: each image was presented for 500 msec, and followed by a blank screen for100 msec. Colours indicate recurring triplets (in red: shoe-fish-airplane; blue: dog-chair-cat). Triplets were created randomly for each subject. Images within the same triplet always appeared in the same order. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

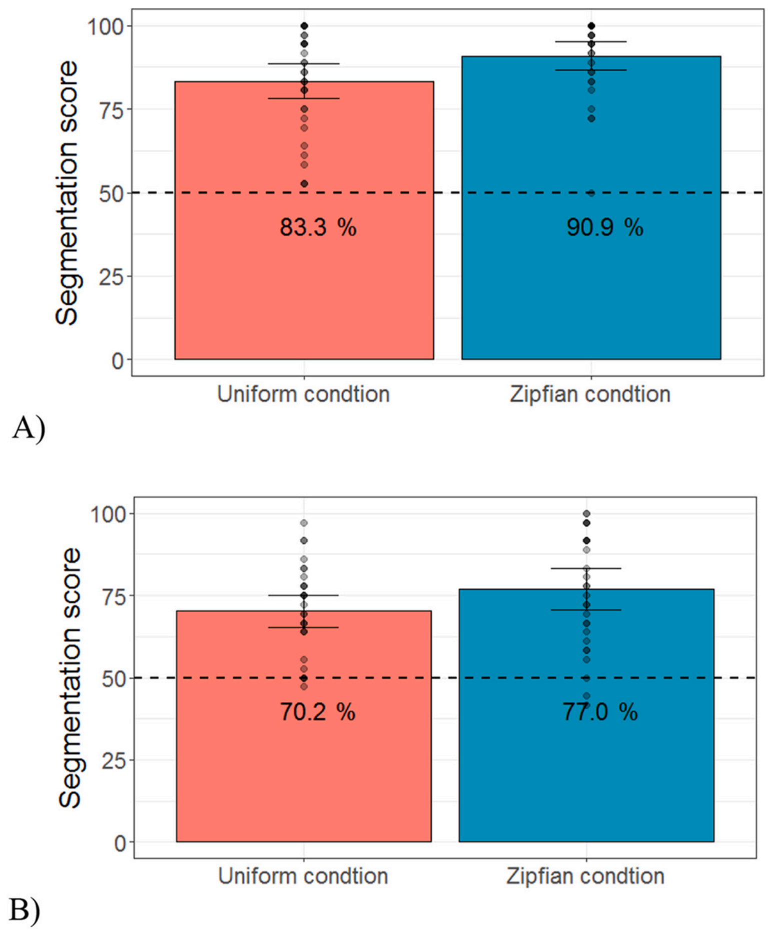


Fig. 3. (A) Adults' and (B) children's segmentation accuracy across conditions. Dashed lines represent chance level. Error bars represents confidence intervals of 95%. Points represents individual scores: darker points indicate more participants with the same score.

Table 2
Mixed-effect regression model comparing adult and child segmentation in the uniform and the Zipfian conditions. Significant variables in bold. Significance obtained using the lmerTest function in R.

	Estimate	Std. Error	z value	*p*-value
(Intercept)	1.452759	0.302901	4.796	<0.001 ***
Zipfian condition	0.990116	0.325019	3.046	<0.01 **
Age group (child)	−1.041327	0.299476	−3.477	<0.001 ***
Triplet's log frequency	0.134362	0.045697	2.940	<0.01 **
Trial number (centered)	−0.019492	0.003967	−4.913	<0.001 ***
Order of appearance (word)	−0.093165	0.082373	−1.131	>0.1
Zipfian condition * Age group (child)	−0.307083	0.441459	−0.696	>0.1

only by better learning of the high frequency triplet, we ran an additional analysis with two changes. First, we excluded the frequent triplet in the Zipfian condition to ensure the advantage held when it was removed. This left 30 trials per subject (instead of 36) in the Zipfian condition (the average triplet frequency was now 10.4 in the Zipfian vs. 24 in the uniform). The second change was to control for the possible effect of "frequent foils": An alternative explanation to our results is that participants in the Zipfian condition used their knowledge of the frequent triplet to rule out foils that contained an image from it. For example, if the frequent triplet contained an image of a shoe, they could rule out any foil that also had a shoe. To control for this, we added a binary variable saying whether the foil in each trial was "frequent" (containing an image from the frequent triplet) or not. Half of the foils for each subject were "frequent foils", and half were not. If the facilitative effect in the Zipfian condition goes beyond learning of the frequent triplet, we should see a significant effect of distribution type after making these two changes. We ran a general linear regression model similar to the previous one, after excluding the frequent triplet and adding "frequent foil" as a fixed effect (see Table 3, Fig. 4). While participants were better at ruling out frequent foils ($\beta = 0.37$, $SE = 0.13$, $p < 0.001$), the effect of distribution remained significant: even though triplet frequency was now lower in the Zipfian distribution, accuracy was still higher than in the uniform distribution ($\beta = 0.68$, $SE = 0.33$,

Table 3
Mixed-effect regression model comparing adult segmentation of low frequency triplets in the uniform and the Zipfian conditions. Significant variables in bold. Significance obtained using the lmerTest function in R.

	Estimate	Std. Error	z value	p-value
(Intercept)	2.07606	0.21725	9.556	<0.001 ***
Zipfian condition	0.67935	0.32533	2.088	<0.05 *
Age group (child)	−1.04195	0.29938	−3.480	<0.001 ***
Is foil frequent	0.36891	0.12820	2.878	<0.001 ***
Trial number (centered)	−0.02067	0.00395	−5.234	<0.001 ***
Order of appearance (word)	−0.10404	0.08233	−1.264	>0.1
Zipfian condition * Age group (child)	−0.30514	0.44129	−0.691	>0.1

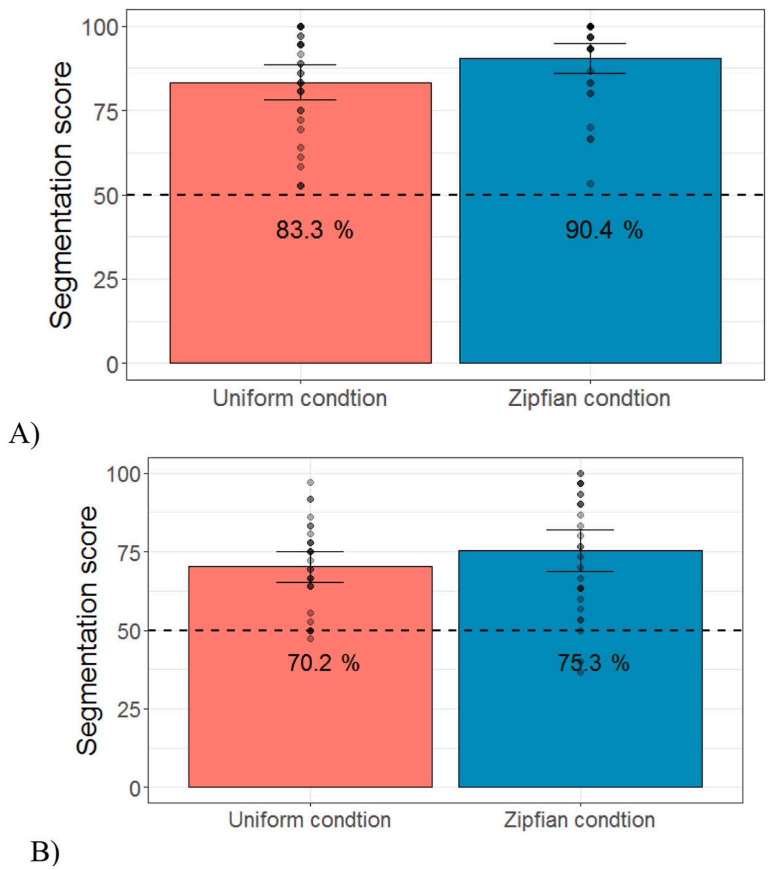


Fig. 4. (A) Adults' and (B) children's segmentation accuracy across conditions, excluding the most frequent triplet in the Zipfian condition (triplet frequency = 24 in the uniform condition, average triplet frequency = 10.4 in the Zipfian condition). Dashed lines represent chance level. Error bars represents confidence intervals of 95%. Points represents individual scores: darker points indicate more participants with the same score.

chi-square(2) = 12.7, $p < 0.05$). Children were less accurate than adults ($\beta = -1.04$, SE = 0.3, $p < 0.001$), but again, the interaction between distribution type and age group was not significant ($\beta = -0.31$, SE = 0.44, $p > 0.1$). We ran an additional model only on the child data to ensure that the effect of distribution type held, and it did (see Appendix C). That is, the facilitative effect holds even after excluding the frequent triplet and controlling for frequent foils.

6. Discussion

The environment we live in, the one whose structure children need to learn, is often non-uniform, with some elements appearing more frequently than others. While such skewed distributions are more predictable than uniform ones, a property which may assist learning, little work to date has examined their impact on learning. Several recent studies found that exposure to skewed distributions facilitates performance across linguistic SL tasks (e.g., Hendrickson & Perfors, 2019; Kurumada et al., 2013; Lavi-Rotbain & Arnon, 2019a, 2019b; Lavi-Rotbain & Arnon, 2020a). Here, we asked whether this effect extends to the visual domain, where, in contrast with the linguistic domain, there is less information about the distribution of units and elements, and where the input is not traditionally thought of as being skewed. Specifically, we wanted to see if visual SL (the detection of visual co-occurrence patterns) will be facilitated in a Zipfian distribution. Though studied extensively, all prior visual SL studies exposed learners to a uniform distribution where each visual unit (pair or triplet) was equally probable. We started by looking at the visual input itself: Using a dataset of objects in infants' visual field, we found that, like individual objects (Clerkin et al., 2017), object pairs and triplets have a right skewed distribution. We then used a standard visual SL paradigm to show that children and adults benefit from this natural skew: both showed better learning of visual triplets when exposed to a Zipfian distribution compared to a uniform one. Importantly, the effect was not driven only by better learning of the most frequent triplet (which is expected to be learned better because of its' higher frequency): The advantage held even when looking only at the lower frequency triplets and controlling for foil frequency. That is, despite appearing fewer times, triplets were learned better in the Zipfian distribution. Even though the relations to be learned were similarly informative (the TPs between images were higher within triplets than between triplets in both distributions), learning was improved when the environment itself was skewed. It is important to note that the documented reliance on TPs for segmentation does not stand in contrast with the facilitative effect of distribution predictability. In the real world, learners are exposed to both cues simultaneously: TPs within words are generally higher than between words, and the words themselves follow a Zipfian distribution making certain words more predictable than others. This means that learners have multiple sources of information to predict what will happen next.

These findings have several theoretical and methodological implications. They document the skewed nature of the visual environment and suggest that the facilitative effect of such distributions is domain-general. They highlight a novel parallel between visual and auditory SL, and between linguistic and non-linguistic SL: Despite being differentially impacted by prior knowledge (e.g., Siegelman et al., 2018) and having different developmental paths (Raviv & Arnon, 2018), the effect of distribution type on learning seems similar across stimulus-type and modality (at least for stimuli whose real-world distribution is skewed). The findings highlight the deep impact of environment structure on learning, and the importance of using the real-world environment as a basis for our lab-based investigations. While the use of uniform distributions is advantageous for isolating particular properties of the input, it comes with the potential cost of underestimating learners' ability. This may be particularly risky when conducting developmental research examining what children (or infants) can or cannot learn. Methodologically, manipulating distribution predictability can open up new ways of changing existing paradigms to better detect learning.

More broadly, the findings suggest that the right skewed distributions found in children's real-world learning environments are beneficial for learning. While the source of such distributions is heavily debated (Piantadosi, 2014), with explanations ranging from purely mathematical (e.g., Chater & Brown, 1999) to more communicatively-oriented (e.g., Ferrer i Cancho & Sole, 2003; Ferrer-i-Cancho, Bentz, & Seguin, 2020), their presence and propensity in the environment may nevertheless have implications for human learning. But what about such distributions facilitates learning? On their own, the current findings are consistent with the impact of prediction on human cognition more generally: an example of better learning in more predictable environments (e.g., Clark, 2013; McClelland et al., 2010). They do not tell us whether there is something uniquely facilitative about Zipfian distributions (compared to other skewed distributions), and what about such distributions facilitates learning. One explanation ties the facilitation to ambiguity reduction and predicts it will not be found in unambiguous learning settings (i.e., when learning already segmented object-label associations, Hendrickson & Perfors, 2019). Under this account, the facilitation is driven by the presence of a highly frequent element, a property shared by many non-Zipfian distributions. Another explanation suggests that Zipfian distributions may confer a unique advantage that has to do with their particular distribution predictability, and the recurring contrast between more and less frequent elements (Lavi-Rotbain & Arnon, 2019b; Lavi-Rotbain & Arnon, 2020a). The graded difference in frequency, which is a hallmark of Zipfian distributions, could facilitate learning by making each higher frequency element an anchor for learning less frequent ones (so that the higher frequency element can assist learning of lower frequency ones across the continuum). This proposal receives preliminary support from corpus analyses showing that different languages have similarly predictable distributions, and experimental findings suggesting the effect of distribution predictability is discontinuous, with certain levels being more optimal for learning (Lavi-Rotbain & Arnon, 2019b). Importantly, much additional work is needed to identify exactly what about such distributions facilitates learning, whether they are more facilitative than other skewed distributions, and whether similar properties of the distribution affect learning across modality.

The current study is a first step in assessing the distribution of object combinations in the real world, and the impact of this distribution on learning the relations between them. While many questions remain open, our findings reveal an effect of distribution type on visual SL, indicating that exposure to skewed distributions facilitates learning across modality and age.

Author contributions

O.L-R. and I.A. conceptualized the study. O.L-R designed the experiments and analyzed the data. O.L-R. and I.A. wrote the manuscript.

Declaration of Competing Interest

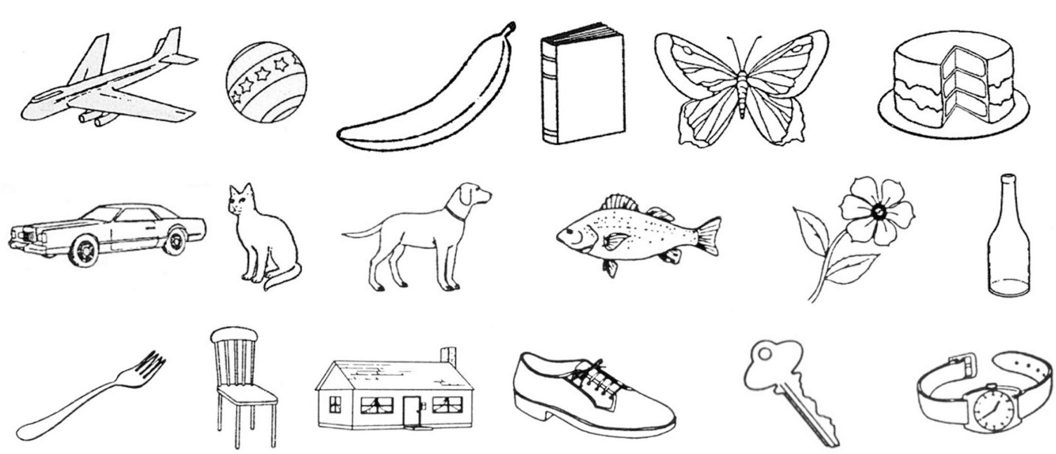
None.

Acknowledgment

We thank Alen Viner for help with analyzing the database of visual objects. We thank Zohar Aizenbud for help with preparing the study. We thank Ephrat Simhon and Shmuel Haikin for their help in collecting the data. We thank Noam Siegelman and Shira Tal for feedback on previous versions of the paper. We thank the Living Lab staff and the Bloomfield Science Museum in Jerusalem, as well as the parents and children who participated. The research was funded by the Israeli Science Foundation grant number 584/16 awarded to the second author.

Appendix A. All images use in Study 2

This is the set of 18 images, taken from (Alario & Ferrand, 1999a):



Appendix B. Segmentation accuracy by age group and triplet frequency

Table B.1 shows mean accuracy of triplets in the Zipfian condition by frequency for each age group separately.

Table B.1
Accuracy in the Zipfian distribution by triplet frequency.

Triplet Frequency	Children	Adults
92	85.6%	93.3%
22	77.8%	93.3%
12	72.2%	91.1%
8	75.0%	89.4%
6	77.2%	86.7%
4	74.4%	91.7%

Appendix C. Children show higher accuracy in the Zipfian distribution

We ran an additional model to ensure that the beneficial effect of the Zipfian distribution on infrequent triplets holds for the children's data alone, and while controlling for age (in months). We compared children's accuracy on all triplets in the uniform condition (36 trials per subject, mean frequency = 24), with their accuracy on the lower frequency triplets in the Zipfian condition (30 trials per subjects, average frequency = 10.4, see Fig. 4B). We used a mixed-effect linear regression model with accuracy on a single trial as our binomial dependent variable, and with the following fixed effects: age (centered), distribution type (uniform condition as baseline); is foil frequent (binary); trial number (centered); and order of appearance in the test. The model had random intercepts for participants as in the previous analyses (see Table C.1). To examine the overall effect of distribution type we used model comparisons. As predicted, the effect of distribution type was significant, even for the lower frequency triplet: Children were more accurate in the Zipfian condition compared to the uniform one, despite the lower frequency of the triplets ($\beta = 0.55$, SE = 0.25, $p < 0.05$; chi-square(1) = 3.93, $p < 0.05$). Older children were more accurate ($\beta = 0.26$, SE = 0.11, $p < 0.05$), in line with previous developmental findings on visual SL (Raviv & Arnon, 2018).

Table C.1
Mixed-effect regression model comparing children segmentation of low frequency triplets. Significant variables in bold. Significance obtained using the lmerTest function in R.

	Estimate	Std. Error	z value	p-value
(Intercept)	2.705078	0.711395	3.802	<0.001 ***
Age (centered)	0.258699	0.105836	2.444	<0.05 *
Zipfian condition	0.518416	0.260848	1.987	<0.05 *
Is foil frequent	0.291557	0.154807	1.883	0.06 .
Trial number (centered)	−0.015037	0.004986	−3.016	<0.01 **
Order of appearance (word)	−0.125789	0.104250	−1.207	>0.1

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104492.

References

Alario, F., & Ferrand, L. (1999a). A set of 400 pictures standardized for French: Norms for name agreement, image agreement. *Behavior Research Methods, Instruments, and Computers, 31*(3), 531–552.
Alario, F., & Ferrand, L. (1999b). 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, and Computers, 31*(3), 531–552. Retrieved from http://www.lpelab.org/userpage/ferrand/PDF/Alario-Ferrand99-BRMIC.pdf%0Apapers3://publication/uuid/FD101C0E-2F70-4468-B766-5DAE78D15364.
Arnon, I. (2020). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods, 52*(1), 68–81. https://doi.org/10.3758/s13428-019-01205-5.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001.
Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me. *Psychological Science, 16*(4), 298–304. https://doi.org/10.1111/j.0956-7976.2005.01531.x.
Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*(1), 127–132. https://doi.org/10.1016/j.cognition.2011.06.010.
Chater, N., & Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition, 69*(3), 17–24. https://doi.org/10.1016/S0010-0277(98)00066-3.
Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477.
Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1711), 20160055. https://doi.org/10.1098/rstb.2016.0055.
Ferrer i Cancho, R., & Sole, R. V. (2003). Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences, 100*(3), 788–791.
Ferrer-i-Cancho, R., Bentz, C., & Seguin, C. (2020). Optimal coding and the origins of Zipfian Laws. *Journal of Quantitative Linguistics, 00*(00), 1–30. https://doi.org/10.1080/09296174.2020.1778387.
Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, 1–87. https://doi.org/10.1037/bul0000210.
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.
Hendrickson, A. T., & Perfors, A. (2019). Cross-situational learning in a. *Zipfian Environment, 189*(May 2017), 11–22. https://doi.org/10.1016/j.cognition.2019.03.005.
James, W. (1890). *The Principles of Psychology* (vol. 2). New York: Henry Holt.
Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). *Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. 83* pp. 4–5).
Kurumada, C., Meylan, S. C., & Frank, M. C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition, 127*(3), 439–453. https://doi.org/10.1016/j.cognition.2013.02.002.
Lavi-Rotbain, O., & Arnon, I. (2019a). Children learn words better in low entropy. In *Proceedings of the 41thth Annual Conference of the Cognitive Science Society*.
Lavi-Rotbain, O., & Arnon, I. (2019b). Low entropy facilitates word segmentation in adult learners. In *Proceedings of the 41th Annual Conference of the Cognitive Science Society*.
Lavi-Rotbain, O., & Arnon, I. (2020a). The learnability consequences of Zipfian distributions: Word Segmentation is Facilitated in More Predictable Distributions. *PsyArXiv*. https://doi.org/10.31234/osf.io/xwgpk.
Lavi-Rotbain, O., & Arnon, I. (2020b). *Zipfian Distributions in Child-Directed Speech* (submitted).
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. https://doi.org/10.1017/S0305000999004006.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences, 14*(8), 348–356. https://doi.org/10.1016/j.tics.2010.06.002.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*(5), 323–351. https://doi.org/10.1080/00107510500052444.

Piantadosi, S. T. (2014). Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic Bulletin & Review, 21*(5), 1112–1130. https://doi.org/10.3758/s13423-014-0585-6.

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*(4), 1–13. https://doi.org/10.1111/desc.12593.

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

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *SSRN*. https://doi.org/10.1146/annurev-psych-122216-011805.

Schuler, K. D., Reeder, P. A., Newport, E. L., & Aslin, R. N. (2017). The effect of Zipfian frequency variations on category formation in adult artificial language learning. *Language Learning and Development, 13*(4), 357–374. https://doi.org/10.1080/15475441.2016.1263571.

Shufaniya, A., & Arnon, I. (2018). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cognitive Science, 42*, 3100–3115. https://doi.org/10.1111/cogs.12692.

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition, 177*(April), 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.

Zipf, G. K. (1949). *Human behavior and the principle of least effort. Human behavior and the principle of least effort*. Oxford, England: Addison-Wesley Press.