

## Review

## Constraints on Statistical Learning Across Species

Chiara Santolin<sup>1,\*</sup> and Jenny R. Saffran<sup>2</sup>

Both human and nonhuman organisms are sensitive to statistical regularities in sensory inputs that support functions including communication, visual processing, and sequence learning. One of the issues faced by comparative research in this field is the lack of a comprehensive theory to explain the relevance of statistical learning across distinct ecological niches. In the current review we interpret cross-species research on statistical learning based on the perceptual and cognitive mechanisms that characterize the human and nonhuman models under investigation. Considering statistical learning as an essential part of the cognitive architecture of an animal will help to uncover the potential ecological functions of this powerful learning process.

## Finding Structures in the Sensory World

A central problem in the study of cognition and development concerns the ways in which organisms compute sensory inputs and discover patterns in the environment. **Statistical learning** (see [Glossary](#)) has been proposed as a key mechanism for extracting regularities distributed across sensory **modalities** and cognitive **domains**, and as a process that constitutes the foundation of further abilities ([1–3] for recent reviews). Over the past 20 years a substantial body of research has explored statistical learning across myriad domains in both human and nonhuman learners.

In the present review we identify constraints on statistical learning: differences in the amount of statistical information acquired or types of computations performed over a structured input by a given organism. We examine similarities and differences between species, and interpret statistical learning findings based on what is known about **domain-general** mechanisms possessed by the species in question.

Researchers in the field of human statistical learning consider these mechanisms to be an essential component of our perceptual and cognitive systems [3]. In humans, statistical learning is constrained in the way in which it operates across modalities and domains [1,4,5]. In addition, data from human infants suggest that statistical learning abilities may emerge at distinct points in development of different sensory modalities. This trajectory is likely explained, at least in part, by the ways in which the perceptual and cognitive skills of infants develop as a function of age and experience.

In nonhuman organisms, it is still unknown how statistical learning relates to other perceptual and cognitive capacities and to **ecological niches**. Studies have largely focused on the types of statistical computations performed by each species, and the extent to which such computations mirror those performed by humans. Less attention has been directed towards the ways in which statistical learning is constrained in nonhuman animals. As in prior work with humans, our goal is to examine statistical learning abilities through the lens of domain-general abilities that differentiate species.

## Trends

The ecological relevance of statistical learning is mostly undefined in the animal kingdom; the majority of cross-species research in this area focuses on revealing human-like abilities rather than discovering the functions of statistical learning in different species.

Perception, memory, and learning guide and constrain statistical learning differently across species, limiting the extent to which organisms can process regularities from sensory inputs.

Cross-species differences can be framed around general-purpose abilities, developmental processes, and learning challenges characterizing the animal models under investigation, and can be interpreted by considering how statistical learning is integrated within the cognitive system of the learner.

<sup>1</sup>Center for Brain and Cognition, Universitat Pompeu Fabra, Carrer Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain  
<sup>2</sup>Waisman Center, University of Wisconsin–Madison, 1500 Highland Avenue, Madison, WI 53705, USA

\*Correspondence: [chiara.santolin@upf.edu](mailto:chiara.santolin@upf.edu) (C. Santolin).

## Who Can Do What?

Statistical learning facilitates the detection of sequential patterns in visual, auditory, and tactile information, as well as of spatial patterns in the visual domain [6–27]. In humans, and some songbird species, statistical learning also supports communicative functions. This process has been implicated in numerous aspects of early language development, including discovering word boundaries, prosodic and phonotactic patterns, syntactic structures, and label–object mappings [28–36]. In a similar vein, vocal learning in songbirds may involve statistical learning processes. For example, juvenile Bengalese and zebra finches learn their own songs by tracking the probability distribution of syllables within songs sung by adult tutors [37–39].

We have chosen to focus this review on three specific statistical learning abilities: tracking sequential statistics, generalizing sequential patterns, and acquiring simple syntactic structures. These three abilities are particularly relevant because they have been studied using comparable behavioral methods across species. Note that there is also a large literature on testing human adults. For the purposes of the current comparative review, however, we focus primarily on human infant studies because the methods are more comparable to those used with nonhuman animals (Box 1).

## Tracking Sequential Patterns

Human neonates detect **frequencies of co-occurrence** of streams of syllables [40] and shapes [41]. The youngest ages for which there is published evidence showing detection of **transitional probabilities** is 4.5 months [42] for visual materials and 7 months for auditory materials [43]. Detecting transitional probabilities is more complex than detecting co-occurrence frequencies because transitional probabilities entail predictive relations among items, whereas frequencies involve simple co-occurrence of elements. Given the studies published to date, it is unclear whether the detection of sequential statistics emerges earlier than is reported in the literature.

Sensitivity to sequential statistics has also been demonstrated in nonhuman species. Zebra finches (*Taeniopygia guttata*) track transitional probabilities to encode sequences of song syllables [44]. When learning songs from adult tutors, juvenile Bengalese finches (*Lonchura striata domestica*) select chunks of notes with greater internal transitional probabilities over groups of notes spanning chunk boundaries [37]. Following mere exposure without reward, newborn domestic chicks (*Gallus gallus*) detect patterns in streams of visual elements [13]. Although it remains unclear exactly which statistics chicks are computing (e.g., transitional probabilities or co-occurrence frequencies), this species does appear to be tuned to distributional information in its postnatal environment.

Among mammals, statistical learning from linguistic materials has been demonstrated in rats (*Rattus norvegicus*), who utilize frequencies of co-occurrence rather than transitional probabilities to track syllables in speech streams [8]. Cotton-top tamarins (*Sanguinus oedipus*) show similar capacities; however, it is unknown whether tamarins detect co-occurrence frequencies or transitional probabilities between syllables [45].

## Generalizing Sequential Patterns

The sequential learning tasks described in the previous section require learners to track specific sequences of elements. Other experimental paradigms assess whether learners can abstract beyond specific sequences, thus requiring **generalization**. Human neonates generalize the structure of triplets of syllables arranged in an ABB pattern (e.g., *ga-ti-ti*, *we-fo-fo*, *la-gu-gu*), discriminating novel ABB sequences from random sequences of the same syllables [46]. A few

## Glossary

**Artificial grammar:** a nonsense language composed of novel words whose order mimics the structures of natural languages.

**Domains:** general areas of cognition such as language, music, object recognition, memory, and attention.

**Domain-general:** refers to processes and abilities that operate across distinct cognitive domains. Statistical learning is often considered to be a domain-general mechanism because it operates across multiple domains including language, music, and visual objects.

**Ecological niche:** a habitat or environment in which a species lives, and behavioral responses emitted to best adapt to that habitat.

**Finite-state grammar:** transitional probabilities between a set of adjacent items (AB)<sup>n</sup>. A unique pair of items generates a linear structure based on the number repetitions required (e.g., ABABABAB).

**Frequency of co-occurrence:** how often two elements composing a pattern appear together. Elements can be syllables, musical notes, or visual shapes based on the sensory modality implementing the pattern.

**Generalization:** learning of a structure independent of the perceptual identity of its components. For example, ABA represents a specific ordering of categories of items (A and B), which can be implemented by any given syllables or shapes; 'cross–circle–cross' and 'triangle–hexagon–triangle' represent instances of the same structure. Generalization requires learners to transfer this structure to novel stimuli (new syllables or shapes).

**Modality:** sensory presentation of the stimuli used in experimental tasks. In the statistical learning literature, the principal modalities are auditory, visual (which can be spatial or temporal), and tactile.

**Nonadjacent dependency:** conditional probability between two items interleaved by at least one additional item (AXB). As for transitional probabilities, A predicts B.

**Phrase structure grammar:** categories of items (usually words such as determiners, nouns, verbs, etc.) that are clustered into

## Box 1. Behavioral Procedures

## (i) Habituation

Infants are exposed to structured visual or auditory sequences until they reach a habituation criterion (diminished looking). Each test trial consists of presentation of a single sequence that is either consistent or inconsistent with the habituation materials. Discrimination is measured as a function of looking time to the consistent versus inconsistent test items.

## (ii) Head-Turn Preference

Infants are exposed to sound sequences for a few minutes in the presence of a neutral visual stimulus. The test items consist of sequences that are either consistent or inconsistent with the exposure materials (e.g., words vs non-words, the same syllables in novel orders). Infants control the duration of each test trial via a head turn in the direction of the auditory stimulus. Discrimination is measured as in habituation. In cotton-top tamarins the orienting response towards the speaker playing the stimuli serves as measure of discrimination.

## (iii) Go/No-Go

Animals are first trained to discriminate a sound associated with the Go response (e.g., pecking on a sensor) from a sound associated with the No-Go response (no pecking) via reinforcement and/or punishment. Test items alternate with training stimuli which are reinforced to avoid response extinction. Subjects categorize test stimuli as consistent with Go or No-Go stimuli experienced during training; discrimination is measured as the proportion of correct responses.

## (iv) Spontaneous Discrimination

Animals are familiarized with artificial languages for hours, and are then tested on strings of sounds consistent versus inconsistent with the familiarization language; discrimination is measured as a function of changes in calling behavior in response to the stimuli. In the newborn chick task, animals are exposed for hours to a structured visual stream of objects. Discrimination between consistent versus inconsistent stimuli is measured by the proportion of time spent near the screen playing the consistent stimulus (Figure 1).

## (v) Operant Conditioning

Animals are trained to select one type of stimuli associated with food (S+); responses toward S− are associated with bland punishment or no-reward. Discrimination is measured as a function of choice of consistent versus inconsistent stimuli with respect to S+. Alternatively, the training phase consists of presentation of stimuli associated with food reinforcement. Test items are words consistent with the language versus non-words, and the behavioral response (e.g., lever-pressing) associated with test stimuli is measured.

subgroups such as noun-phrases and verb-phrases. Subphrases can be nested into other subphrases to confer hierarchical organization to a sentence. Which word pertains to which category, and the position of subphrases within a sentence, is determined by statistical dependencies (an example of rudimentary phrase structures in the literature:  $A^nB^n$ ; AAABBBBBB).

**Statistical learning:** learning mechanism enabling detection of regularity (e.g., co-occurrence frequencies, transitional probabilities, nonadjacent dependencies, etc.) in sensory inputs.

**Transitional probability:** a form of conditional probability defined by the formula: probability of B | A = frequency of AB/frequency of A. This conditional statistic is sensitive to the order with which one item predicts the following one within a pair (AB).

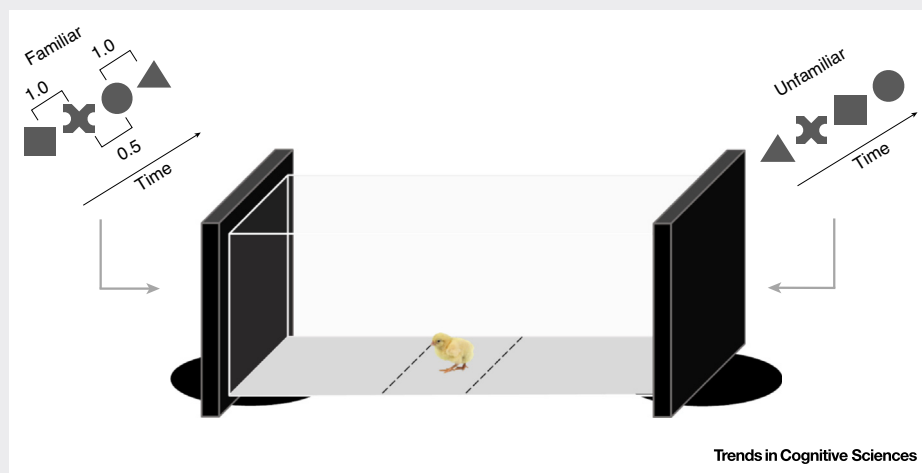


Figure 1. Apparatus and Sample Stimuli Used to Investigate the Detection of Statistical Patterns in Chicks [13]. In the familiar sequence, the shapes are structured into pairs such that the first shape in a pair is always followed by the same second shape. In the unfamiliar sequence, the same shapes are presented but in random order.

months later, infants' generalization abilities become more robust. After familiarization with a pattern such as ABB, 7-month-old infants discriminate between novel syllables arrayed in this familiar structure from the same syllables in a new structure (e.g., *ga-ti-ti* vs *ga-ti-ga* [47]).

Infants perform similar computations in visual and auditory nonlinguistic domains, although with some constraints. It has been suggested that infants perform better if previously exposed to the same regularities implemented by speech ([48]; for counterexamples, see [49]), or, more broadly, communicative signals [50]. An alternative hypothesis suggests that successful learning is impacted by familiarity and/or ease of processing. For example, infants are more apt to generalize sequences of animals or upright faces than geometric forms or inverted faces [9,21,51]. Other perceptual and cognitive factors appear to constrain this form of learning in infants, including the presence of immediate repetitions of individual elements [9,21,52].

Nonhuman animals also generalize in sequential learning tasks, but with notable limitations. Zebra finches generalize based on the perceptual similarity between training and test materials rather than syllable order. In fact, in the absence of acoustic properties shared between familiar and novel streams, finches fail to generalize at all [53–55]. Similarly to human infants, finches seem to privilege patterns containing adjacent repetitions such as ABB and AAB [54]. Newly hatched chicks demonstrate robust generalization given training on triplets of visual objects arranged according to ABA, AAB, ABB, and BAA patterns, regardless of the presence of immediate repetitions [26]. Among mammals, rats trained on ABA sequences instantiated by strings of tones generalize to novel stimuli [56]. Similar results have been obtained with consonant–vowel alternations implementing ABB patterns compared to random sequences [57,58]. Rhesus macaques (*Macaca mulatta*) discriminate novel AAB versus ABB strings implemented by their own calls [59]; a recent extension of this study involved another primate species, the cotton-top tamarin, which can generalize structures presented in both speech and musical tones [60].

### Acquiring Simple Syntactic Structures

By 12 months of age, human infants can learn rudimentary syntactic structures in laboratory tasks. Using **artificial grammar** learning paradigms, infants are first familiarized with simple miniature grammars that generate sets of sentences. Infants subsequently discriminate grammatical versus ungrammatical strings containing either violations of internal syllable pairs or violations at the edges of the grammar [32]. In a similar paradigm, 12-month-olds learned **phrase structures grammars** that mimicked the basic statistical patterns of natural languages (e.g., a determiner, such as 'the' or 'a', requires a noun somewhere within a sentence [34]). In this latter study, nonsense words from an artificial language were clustered in categories (e.g., determiners) whose presence depended on the presence of other categories (e.g., nouns). After exposure to the language, infants distinguished grammatical versus ungrammatical test strings which violated predictive dependencies between word categories. This task also required generalization of grammatical knowledge beyond the trained sentences.

The ability of songbirds to learn and generalize syntactic structures is somewhat more restricted. European starlings (*Sturnus vulgaris*) learn structures as complex as **finite-state grammar** (e.g.,  $AB^n$ ) and hierarchical grammar (e.g.,  $A^nB^n$ ) formed by their own song syllables [61]. Finite-state grammars generate strings of items repeated linearly (e.g., ABABAB) whereas hierarchical grammars comprise categories of items grouped into substrings that are nested into other substrings (e.g., phrase structure, AABBBB). However, at least in starlings, generalization to novel instances of the grammars (e.g., CDCDCD) is driven by acoustic similarity between training and test syllables, rather than by detection of the underlying patterns [62].

When the computations require transfer of structured information to novel inputs, these species exhibit limited capacities. Bengalese finches (*Lonchura striata domestica*) demonstrate more advanced generalization skills, taking advantage of statistical (predictive) dependencies between categories of song syllables, and generalize to strings composed of novel syllables [10]. This evidence is consistent with the previously described results from 12-month-old human infants, pointing to common processing of basic syntactic patterns between the two species.

Among nonhuman primates, tamarins learn finite-state and phrase-structure grammars [34,63]. However, in the latter case, monkeys failed to extract the structure when it required generalization to novel sentences, unlike human infants, suggesting that the operation is limited to the specific stimuli with which they were familiarized. In a series of artificial grammar learning tasks involving phrase-structure grammars (similar to [34]), marmosets (*Callithrix jacchus*) primarily encoded regularities involving the beginning of sentences, whereas macaques could also track violations throughout the strings [11]. Macaques have also been compared to human adults in tasks investigating learning of **nonadjacent dependencies**. In line with some findings on nonhuman primates [63,64], but not others [65], macaques exhibited no sensitivity to regularities involving nonadjacent syllables, only responding to violations of contiguous syllables. Humans performed better than macaques overall, detecting nonadjacent patterns [66].

### Species Comparisons and Constraints

Despite many similarities, human infants and nonhuman animals diverge in the facility with which they perform various statistical learning tasks. Our interpretation of these cross-species differences considers perception and cognition, focusing on the extent to which these general abilities constrain statistical learning. A similar approach has been taken to explain early learning of grammar with respect to similarities between human and nonhuman performance. According to this perspective, grammar acquisition is supported by specialized perceptual and memory systems that are possibly shared with other species [67]. In the present review article we examine and interpret cross-species differences in comparable tasks based on domain-general abilities possessed by the species in question. Considering the creature under investigation *tout court* has the potential to provide a window into the ecological functions of statistical learning. To this end, we will examine statistical learning cross-species comparisons through the lens of three aspects of perception and cognition: vocal learning, perceptual processing, and memory.

#### Relationship between Vocal Learning and Statistical Learning

Vocal learning requires animals to modify the acoustic structure of the vocalizations of their own species and produce novel patterns of sounds. Through this mechanism, the young of some species acquire fundamental communicative signals that facilitate social interactions with conspecifics. Vocal learners include passerine songbirds, seals, cetaceans, and some bats. However, there are substantial differences in the way this process unfolds across species. Some species acquire a single novel vocalization as juveniles, whereas others learn new sounds even in adulthood [68,69]. Differences also include the structure of the learned sound patterns. The songs of some species provide more structural variability (e.g., European starlings), whereas others consist of fixed sequences of notes (e.g., zebra finches) [70,71]. Differences in vocal learning across species may predict the complexity of the patterns animals can learn and generalize in laboratory tasks [55]. In particular, species whose songs are syntactically structured and composed of a wide range of notes should be better at performing complex computations requiring, for instance, generalization of a given pattern to novel exemplars.

Consistent with this hypothesis, recent findings comparing two avian species in an artificial grammar learning task reveal differences that are consistent with general aspects (syntactic-like organization, phonological variability) of the vocal behavior of birds [55]. Zebra finches learn regularities such as ABA and AAB presented as short sequences of song syllables, but fail to generalize to structured strings formed by novel song syllables. However, the budgerigar (common parakeet, *Melopsittacus undulatus*) goes beyond item-specific information and generalizes these patterns, recognizing them even when they are implemented in novel sounds. This difference in performance can be explained in light of what is known about vocal learning in these species. The vocal repertoire of zebra finches includes highly stereotyped songs formed by rigid syllable sequences, resulting in repetitive, linear patterns [70–73]. Budgerigars are open-ended vocal learners, with high vocal plasticity, and their songs show greater phonological variability and syntactic-like organization [74,75]. We hypothesize that the perceptual and learning skills possessed by the finches may not be suited to performing computations more complex than rote memorization of specific syllable order. By contrast, the vocal learning abilities of budgerigars suggest the presence of learning mechanisms that can detect the underlying structure of sound sequences, later recognizing that structure when implemented by novel syllables (e.g., generalization). This hypothesis is supported by findings showing that budgerigars possess generally superior memory for acoustic complex stimuli [76]. The directionality of this relationship remains unknown. In line with theorizing about human language, we hypothesize that the learning abilities themselves have shaped the song structure of these species [4,77].

#### Visual Statistical Learning in Neonates

The development of perceptual processing also provides insights into the learning outcomes observed across species. Precocial and altricial species differ in the ontogeny of a range of physiological and behavioral functions. Precocial animals are biologically mature from hatching or birth, and are generally independent from parental care. For instance, superprecocial animals such as megapode birds (e.g., Australian brushturkey, *Alectura lathamii*) leave the nest shortly after hatching, and display fully developed brain structures and motor behavior, being able to fly few hours after hatching [78–80]. As a consequence, the early perceptual processing performed by precocial organisms exceeds that of altricial species (see also [81]). These ontogenetic differences may be linked to the statistical learning skills present in a given species at birth. It is likely the case that altricial animals would have to deal with biological limitations (e.g., neural development, sensory and perceptual processing) that constrain what can be learned at the outset of postnatal life.

From this perspective, consider the comparison between human neonates – an altricial species with an extended developmental timeline – and newly hatched domestic chicks, a precocial bird species. In a visual statistical learning study, newly hatched chicks discriminated structured from random streams of four and six shapes [13]. Human neonates, however, succeeded only with streams of four shapes, failing with six-shape streams [41]. Limited perceptual abilities are likely to constrain visual statistical learning in humans at birth, whose visual system is immature (e.g., severely reduced acuity [82]), generally limiting visual learning [83]. Such restrictions are typical of the altricial human primate, whose offspring stay immature longer than other mammals [84,85]. Unlike humans, chicks hatch in an advanced stage of development, with completely developed visual pathways from the very first days of postnatal life [86–88]. In chicks, vision is the predominant sensory modality, allowing full processing of complex visual stimuli immediately after hatching. Compared to chicks, reduced visual skills in human infants appear to constrain statistical learning, limiting the extent to which infants can process statistics over particular inputs. Indeed, a classic theory of perceptual development hypothesizes that



limitations on the amount of information computed – owing to the protracted maturation of some sensory systems relative to others – actually facilitate human perceptual and cognitive development [89].

Statistical learning is likely to guide different functions in humans and chicks. In humans, early learning of regularities plays a key role in domains other than vision, especially language processing. In chicks, however, extracting regularities might be linked with early social learning, a fundamental capacity that allows newly hatched chicks to recognize the mother hen and siblings. This process, filial imprinting, occurs immediately after birth via learning of invariant visual features of their social companions, such as plumage color patterns as well as beak and head shape [86,90]. We hypothesize that statistical learning works in tandem with filial imprinting in chicks, and leads to an integrated representation of relevant social objects that will allow further recognition and identification. Consistent with our hypothesis, a recent study shows that imprinting in chicks promotes the learning of multimodal regularities such as XX versus XY implemented by sound–shape pairs as well as generalization to novel instances ([91], see also [92]). Being precocial requires chicks to process salient visual stimuli at birth, whereas altricial humans do not have to rely on early visual learning to identify relevant social inputs. In this view, statistical learning abilities are influenced by the ontogeny of the species, which determines the functions supported by the learning process (i.e., language acquisition, recognition of social objects, etc.).

### Memory and Statistical Learning

Memory is a fundamental component of statistical learning. Learners must keep track of sequences of elements that rarely persist over time, and temporarily hold information necessary for future processing (i.e., working memory [93]). According to the memory-based framework presented in [94], regularities extracted from structured input lead to representations in long-term memory even after short exposure (see also [95]). Memory traces then become the foundation for subsequent learning operations, driving detection of further patterns, and integrating stored information to find common regularities across exemplars [96,97]. Computations such as extraction, storing, and integration seem to be fundamental for both acquiring sequential regularities and generalizing to novel exemplars. This framework also points to the way in which learners retain and retrieve information as a constraint on statistical learning, suggesting that sensitivity to patterns in the sensory input is shaped by the memory skills of learners (e.g., storing, access). Following this path, we hypothesize that species with reduced memory skills would be sensitive to a restricted set of structured information from sequentially presented inputs compared to species with enhanced storing and retention abilities.

Among nonhuman primates, marmosets and macaques have been directly compared in artificial grammar learning experiments to test learning of sequential syntactic patterns. When presented with strings violating the familiar grammar at multiple locations, marmosets, unlike macaques, detect only violations at the very beginning of the grammar [11]. One explanation for this pattern of performance points to cross-species differences in memory skills. These species belong to separate groups of the primate order, with distinct physical features and cognitive abilities. Macaques are equipped with superior general learning and memory capacities, outperforming marmosets in basic learning tasks such as the discrimination of learned stimuli [98,99]. In particular, some species of the *Callitrichidae* family, including the common marmoset, seem to use spatial memory while foraging, concurrently tracking several object locations containing food [100–102]. It is likely that such differences affect other computations that closely involve learning, such as statistical processing of grammatical sequences. In this view,

marmosets developed better memory skills in the spatial modality than in the temporal modality, thus performing worse on tasks requiring retention of sequential stimuli.

The general learning systems of nonhuman primates may not be tailored to acquire linguistic structures as complex as hierarchical grammars (see also [103]), but do support learning of sequential patterns in other domains such as tool-use [104,105], motor actions [106,107], and social communication [108] which require the processing of events occurring over time and play a prominent role in the behavioral repertoires of primates. It is possible that the use of linguistic materials in studies investigating sequential statistical learning has hindered the discovery of more advanced abilities in nonhuman primates (see Concluding Remarks). Experiments using ‘grammars’ created not from words but from sequences of tools or actions might reveal statistical learning abilities that are not observed with linguistic materials. More generally, interactions between statistical learning abilities and the elements over which learning occurs (e.g., speech syllables vs motor actions) point to important constraints on statistical learning, and may help to explain divergent learning outcomes across species.

### Concluding Remarks

In this review we have framed cross-species differences in statistical learning in the context of perceptual and cognitive mechanisms that vary across organisms. It is clearly the case that reduced perceptual abilities in a given modality lead to limited learning of structured inputs, as in the case of human neonates. In a similar vein, reduced memory capacity affects regularities detected from sequential inputs, confining learning to a restricted portion of the information available in the input.

According to this perspective, we might expect similar computations in species that share cognitive functions supported by statistical learning. For instance, statistical learning is likely to drive vocal learning in organisms that must learn to produce structured vocalizations [71,109,110]. Similarly to human infants acquiring languages, juveniles of some songbird species track the statistical properties of songs, and recombine statistically coherent patterns into new songs [37,38]. These observations suggest that both acoustic features and the structural organization of the songs are acquired during vocal learning, making this process similar to the beginning of human language acquisition. It is thus not surprising that some bird species and humans exhibit the most advanced statistical learning of sound sequences.

Vocal learning is not limited to birds and humans. Bats, seals, cetaceans, and elephants meet the requirements for vocal learners: they can learn novel vocalizations and imitate the sounds of other species. Among cetaceans, bottlenose dolphins (*Tursiops truncatus*) possess impressive vocal learning and mimicry abilities: they learn to produce new vocalizations, have sophisticated communication systems composed of extended call repertoires, and generally possess remarkable cognitive skills [111–113]. For these reasons we would expect dolphins to show similar statistical learning capacities to human neonates and songbirds [114]. In a similar vein, we would expect non-vocal learners to show restricted abilities in artificial grammar learning tasks. For example, domestic chickens are limited to short repetitive calls [115,116] whose usage is acquired following social experience with conspecifics [110]. Given that there is no vocal learning in this species, we predict that young chicks would fail to track probabilistic structures from auditory stimuli, performing worse than vocal learners.

Statistical learning may support communicative functions even in the absence of vocal learning. Fieldwork suggests that free-ranging putty-nosed monkeys (*Cercopithecus nictitans*) can track co-occurrences between specific combinations of natural calls and the presence of an eagle,

### Outstanding Questions

What is the ecological function of statistical learning in distinct species? What does tracking frequency or probability distributions allow animals to do in real life?

Do cross-species findings imply the presence of a unique statistical learning mechanism that is shared between organisms? Or does the fact that statistical learning leads to different outcomes in different species imply the presence of multiple separate mechanisms?

Is statistical learning domain-general even within a given animal species? For example, is statistical learning involved in domains other than vocal learning in songbirds, as it is in humans?

What is the directionality of the relationship between differences in statistical learning and differences in general perceptual/cognitive abilities across species? For example, do differences in birdsong structure result from different learning mechanisms across species, or are differences in learning mechanisms due to different song structures across species?



eventually emitting vocalizations to signal the predator presence to the rest of the group [117]. Black-fronted titi monkeys (*Callicebus nigrifrons*) produce call sequences whose combination and type vary based on the context (e.g., aerial vs terrestrial predators [118,119]). Female baboons (*Papio cynocephalus ursinus*) appear to possess a combinatorial system of communication reflecting the complex hierarchy of their social group [108]. For example, female baboons notice inconsistent call sequences with respect to the rank of the caller, recognizing when a dominant female emits unusual vocalizations when interacting with subordinates [120].

Statistical learning abilities also impact on non-communicative domains across species. Guinea baboons (*Papio papio*) demonstrated learning of hierarchical structures from visual stimuli [121] and orthographic patterns [23]. Baboons learned statistical relations between printed letters and their positions within a word, distinguishing English words and nonwords with remarkable accuracy. The main theoretical implication of this work is that orthographic processing, an essential computation in reading, might be rooted into more general statistical learning mechanisms shared across baboons and humans, and appears to be constrained by domain-general abilities necessary to discriminate visual objects (e.g., detecting feature combinations).

Further consideration of the role of the perceptual and cognitive mechanisms supporting statistical learning will help to clarify how this process unfolds in different ecological niches (see Outstanding Questions). In human infants, the evidence suggests that statistical learning abilities have impacted on the types of structures that are observed in human languages [4,33,34]. It is also the case that what human infants have already learned impacts on the types of structures that they subsequently learn most readily [96,122]. Both of these directions of effects may influence statistical learning in non-human animals as well. Interpreting cross-species findings in this light links our approach to perspectives on human learning that considers statistical learning as a core component of cognitive systems, rather than as an independent computational process [3]. This view is also aligned with modern neurocomputational theories which treat brains as sophisticated prediction machines [123], internalizing probabilistic models of the environment to anticipate the sensory stream and generate inferences [124,125]. Shifting the focus of comparative research to the systems within which learning is embedded will allow us to develop a much deeper understanding of the ecological relevance of statistical learning.

### Acknowledgments

We are grateful to Alberto Testolin and Erica Wojcik for comments on earlier versions of this manuscript. Preparation of the manuscript was supported by a fellowship from the Foundation Marica De Vincenzi ONLUS in association with the Department of Psychology and Cognitive Science of the University of Trento to C.S., and by a grant from the National Institutes of Health (R37HD037466) to J.S.

### References

1. Frost, R. *et al.* (2015) Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* 19, 117–125
2. Dehaene, S. *et al.* (2015) The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* 88, 2–19
3. Armstrong, B.C. *et al.* (2017) The long road of statistical learning research: past, present and future. *Philos. Trans. R. Soc. B* 372, 20160047
4. Saffran, J.R. (2002) Constraints on statistical language learning. *J. Mem. Lang.* 47, 172–196
5. Krogh, L. *et al.* (2013) Statistical learning across development: flexible yet constrained. *Front. Psychol.* 3, 598
6. Saffran, J.R. *et al.* (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52
7. Conway, C.M. and Christiansen, M.H. (2001) Sequential learning in non-human primates. *Trends Cogn. Sci.* 5, 539–546
8. Toro, J.M. and Trobalón, J.B. (2005) Statistical computations over a speech stream in a rodent. *Atten. Percept. Psychophys.* 67, 867–875
9. Johnson, S.P. *et al.* (2009) Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy* 14, 2–18
10. Abe, K. and Watanabe, D. (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. Neurosci.* 14, 1067–1074

11. Wilson, B. *et al.* (2013) Auditory artificial grammar learning in macaque and marmoset monkeys. *J. Neurosci.* 33, 18825–18835
12. Slone, L.K. and Johnson, S.P. (2015) Infants' statistical learning: 2- and 5-month-olds' segmentation of continuous visual sequences. *J. Exp. Child Psychol.* 133, 47–56
13. Santolin, C. *et al.* (2016) Unsupervised statistical learning in newly hatched chicks. *Curr. Biol.* 26, R1218–R1220
14. Tummeltshammer, K. *et al.* (2016) Across space and time: infants learn from backward and forward visual statistics. *Dev. Sci.* 20, e12474
15. Baldwin, D.A. *et al.* (2001) Infants parse dynamic action. *Child Dev.* 72, 708–717
16. Kirkham, N.Z. *et al.* (2002) Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83, B35–B42
17. Fiser, J. and Aslin, R.N. (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504
18. Fiser, J. and Aslin, R.N. (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 458–467
19. Fiser, J. and Aslin, R.N. (2002) Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15822–15826
20. Fiser, J. and Aslin, R.N. (2005) Encoding multielement scenes: statistical learning of visual feature hierarchies. *J. Exp. Psychol. Gen.* 134, 521–537
21. Saffran, J.R. *et al.* (2007) Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition* 105, 669–680
22. Kirkham, N.Z. *et al.* (2007) Location, location, location: development of spatiotemporal sequence learning in infancy. *Child Dev.* 78, 1559–1571
23. Grainger, J. *et al.* (2012) Orthographic processing in baboons (*Papio papio*). *Science* 336, 245–248
24. Stobbe, N. *et al.* (2012) Visual artificial grammar learning: comparative research on humans, kea (*Nestor notabilis*) and pigeons (*Columba livia*). *Philos. Trans. R. Soc. B* 367, 1995–2006
25. Minier, L. *et al.* (2015) The temporal dynamics of regularity extraction in non-human primates. *Cogn. Sci.* 40, 1019–1030
26. Santolin, C. *et al.* (2016) Generalization of visual regularities in newly hatched chicks (*Gallus gallus*). *Anim. Cogn.* 19, 1007–1017
27. Scarf, D. *et al.* (2016) Orthographic processing in pigeons (*Columba livia*). *Proc. Natl. Acad. Sci. U. S. A.* 113, 11272–11276
28. Saffran, J.R. *et al.* (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926–1928
29. Aslin, R.N. *et al.* (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9, 321–324
30. Mattys, S.L. *et al.* (1999) Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* 38, 465–494
31. Mattys, S.L. and Jusczyk, P.W. (2001) Phonotactic cues for segmentation of fluent speech by infants. *Cognition* 78, 91–121
32. Gomez, R.L. and Gerken, L. (1999) Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135
33. Saffran, J.R. (2001) The use of predictive dependencies in language learning. *J. Mem. Lang.* 44, 493–515
34. Saffran, J.R. *et al.* (2008) Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 107, 479–500
35. Graf-Estes, K. *et al.* (2007) Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol. Sci.* 18, 254–260
36. Clerkin, E.M. *et al.* (2017) Real-world visual statistics and infants' first-learned object names. *Philos. Trans. R. Soc. B* 372, 20160055
37. Takahasi, M. *et al.* (2010) Statistical and prosodic cues for song segmentation learning by Bengalese finches (*Lonchura striata* var. domestica). *Ethology* 116, 481–489
38. Menyhart, O. *et al.* (2015) Juvenile zebra finches learn the underlying structural regularities of their fathers' song. *Front. Psychol.* 6, 1–12
39. Fehér, O. *et al.* (2017) Statistical learning in songbirds: from self-tutoring to song culture. *Philos. Trans. R. Soc. B* 372, 20160053
40. Teinonen, T. *et al.* (2009) Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10, 21
41. Bulf, H. *et al.* (2011) Visual statistical learning in the newborn infant. *Cognition* 121, 127–132
42. Marcovitch, S. and Lewkowicz, D.J. (2009) Sequence learning in infancy: the independent contributions of conditional probability and pair frequency information. *Dev. Sci.* 12, 1020–1025
43. Thiessen, E.D. *et al.* (2005) Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71
44. Chen, J. and ten Cate, C. (2015) Zebra finches can use positional and transitional cues to distinguish vocal element strings. *Behav. Processes* 117, 29–34
45. Hauser, M.D. *et al.* (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78, B53–B64
46. Gervain, J. *et al.* (2008) The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14222–14227
47. Marcus, G.F. *et al.* (1999) Rule learning by seven-month-old infants. *Science* 283, 77–80
48. Marcus, G.F. *et al.* (2007) Infant rule learning facilitated by speech. *Psychol. Sci.* 18, 387–391
49. Dawson, C. and Gerken, L. (2009) From domain-general to domain-sensitivity: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. *Cognition* 111, 378–382
50. Ferguson, B. and Lew-Williams, C. (2016) Communicative signals support abstract rule learning by 7-month-old infants. *Sci. Rep.* 6, 25434
51. Bulf, H. *et al.* (2015) Many faces, one rule: the role of perceptual expertise in infants' sequential rule learning. *Front. Psychol.* 6, 1595
52. Thiessen, E.D. (2012) Effects of inter- and intra-modal redundancy on infants' rule learning. *Lang. Learn. Dev.* 8, 197–214
53. Chen, J. *et al.* (2015) Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Anim. Cogn.* 18, 151–164
54. van Heijningen, C.A. *et al.* (2013) Rule learning by zebra finches in an artificial grammar learning task: which rule? *Anim. Cogn.* 16, 165–175
55. Spierings, M.J. and ten Cate, C. (2016) Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proc. Natl. Acad. Sci. U. S. A.* 113, E3977–E3984
56. Murphy, R.A. *et al.* (2008) Rule learning by rats. *Science* 319, 1849–1851
57. de La Mora, D. and Toro, J.M. (2013) Rule learning over consonants and vowels in a non-human animal. *Cognition* 126, 307–312
58. Crespo-Bojorque, P. and Toro, J.M. (2015) The use of interval ratios in consonance perception by rats (*Rattus norvegicus*) and humans (*Homo sapiens*). *J. Comp. Psychol.* 129, 42
59. Hauser, M.D. and Glynn, D. (2009) Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations? *J. Comp. Psychol.* 123, 161
60. Neiwirth, J.J. *et al.* (2017) Artificial grammar learning in tamarins (*Saguinus oedipus*) in varying stimulus contexts. *J. Comp. Psychol.* 131, 128
61. Gentner, T.Q. *et al.* (2006) Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207

62. Comins, J.A. and Gentner, T.Q. (2013) Perceptual categories enable pattern generalization in songbirds. *Cognition* 128, 113–118
63. Fitch, W.T. and Hauser, M.D. (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377–380
64. Newport, E.L. *et al.* (2004) Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn. Psychol.* 49, 85–117
65. Ravignani, A. *et al.* (2013) Action at a distance: dependency sensitivity in a New World primate. *Biol. Lett.* 9, 20130582
66. Wilson, B. *et al.* (2015) Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *Eur. J. Neurosci.* 41, 568–578
67. Endress, A.D. *et al.* (2009) Perceptual and memory constraints on language acquisition. *Trends Cogn. Sci.* 13, 348–353
68. Catchpole, C.K. and Slater, P.J.B., eds (1995) *Bird Song: Biological Themes and Variations*, Cambridge University Press
69. Okanoya, K. (2004) The Bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann. N. Y. Acad. Sci.* 1016, 724–735
70. Honda, E. and Okanoya, K. (1999) Acoustical and syntactical comparisons between songs of the white-backed Munia (*Lonchura striata*) and its domesticated strain, the Bengalese finch (*Lonchura striata* var. domestica). *Zool. Sci.* 16, 319–326
71. Berwick, R.C. *et al.* (2011) Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* 15, 113–121
72. Price, P.H. (1979) Developmental determinants of structure in zebra finch song. *J. Comp. Physiol. Psychol.* 93, 260
73. Scharff, C. and Nottebohm, F. (1991) Study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* 11, 2896–2913
74. Farabaugh, S.M. *et al.* (1992) Analysis of warble song of the budgerigar *Melopsittacus undulatus*. *Bioacoustics* 4, 111–130
75. Dooling, R.J. *et al.* (1987) Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*). II. Vocal signals. *J. Comp. Psychol.* 101, 367
76. Dent, M.L. *et al.* (2008) Species differences in the identification of acoustic stimuli by birds. *Behav. Proc.* 77, 184–190
77. Christiansen, M.H. and Chater, N. (2008) Language as shaped by the brain. *Behav. Brain Sci.* 31, 489–509
78. Starck, J.M. and Ricklefs, R.E. (1998) Patterns of development: the altricial-precocial spectrum. *Oxf. Ornithol. Series* 8, 3–30
79. Andrew, R.J. (ed.) (1991) *Neural and Behavioural Plasticity: The Use of the Chick as a Model*, Oxford University Press
80. Rose, S. (2000) God's organism? The chick as a model system for memory studies. *Learn. Mem.* 7, 1–17
81. Sloman, A. and Chappell, J. (2005) The altricial-precocial spectrum for robots. *Proceedings International Joint Conference on Artificial Intelligence* 19, 1187–1192
82. Braddick, O. and Atkinson, J. (2011) Development of human visual function. *Vision Res.* 51, 1588–1609
83. Johnson, M.H. and De Haan, M. (2015) *Developmental Cognitive Neuroscience: An Introduction*, John Wiley & Sons
84. Zeveloff, S.I. and Boyce, M.S. (1982) Why human neonates are so altricial. *Am. Nat.* 120, 537–542
85. Montagu, M.F. (1960) Time, morphology, and neoteny in the evolution of man. In *An Introduction to Physical Anthropology*, (3rd edn), pp. 295–316, Charles C. Thomas
86. Lorenz, K.Z. (1937) The companion in the bird's world. *Auk* 54, 245–273
87. Schmid, K.L. and Wildsoet, C.F. (1998) Assessment of visual acuity and contrast sensitivity in the chick using an optokinetic nystagmus paradigm. *Vision Res.* 38, 2629–2634
88. Deng, C. and Rogers, L.J. (1998) Bilaterally projecting neurons in the two visual pathways of chicks. *Brain Res.* 794, 281–290
89. Turkewitz, G. and Kenny, P.A. (1982) Limitations on input as a basis for neural organization and perceptual development: a preliminary theoretical statement. *Dev. Psychobiol.* 15, 357–368
90. McCabe, B.J. (2013) Imprinting. *Wiley Interdisciplinary Reviews. Cognitive Sc.* 4, 375–390
91. Versace, E. *et al.* (2017) Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Anim. Cogn.* 20, 521–529
92. Wood, J.N. (2013) Newborn chickens generate invariant object representations at the onset of visual object experience. *Proc. Natl. Acad. Sci.* 110, 14000–14005
93. Awh, E. and Jonides, J. (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5, 119–126
94. Thiessen, E.D. *et al.* (2013) The extraction and integration framework: a two-process account of statistical learning. *Psychol. Bull.* 139, 792–814
95. Kim, R. *et al.* (2009) Testing assumptions of statistical learning: is it long-term and implicit? *Neurosci. Lett.* 461, 145–149
96. Thiessen, E.D. and Saffran, J.R. (2007) Learning to learn: infants' acquisition of stress-based strategies for word segmentation. *Lang. Learn. Dev.* 3, 73–100
97. McClelland, J.L. and Rumelhart, D.E. (1985) Distributed memory and the representation of general and specific information. *J. Exp. Psychol. Gen.* 114, 159–197
98. Miles, R.C. (1957) Delayed-response learning in the marmoset and the macaque. *J. Comp. Physiol. Psychol.* 50, 352–355
99. Miles, R.C. and Meyer, D.C. (1956) Learning sets in marmosets. *J. Comp. Physiol. Psychol.* 49, 212–222
100. Menzel, E.W., Jr *et al.* (1985) Social foraging in marmoset monkeys and the question of intelligence. *Proc. Biol. Sci.* 308, 145–158
101. Garber, P.A. (1989) Role of spatial memory in primate foraging patterns: *Saguinus mystax* and *Saguinus fuscicollis*. *Am. J. Primatol.* 19, 203–216
102. MacDonald, S.E. *et al.* (1994) Marmoset (*Callithrix jacchus jacchus*) spatial memory in a foraging task: win-stay versus win-shift strategies. *J. Comp. Psychol.* 108, 328
103. Fitch, W.T. (2000) The evolution of speech: a comparative review. *Trends Cogn. Sci.* 4, 258–267
104. Greenfield, P.M. (1991) Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* 14, 531–595
105. Piñon, D. and Greenfield, P.M. (1994) Does everybody do it? Hierarchically organized sequential activity in robots, birds and monkeys. *Behav. Brain Sci.* 17, 361–365
106. Jordan, M.I. and Rosenbaum, D.A. (1990) Action. In *Foundations of Cognitive Science* (Posner, M.I., ed.), pp. 727–767, MIT Press
107. Dawkins, R. (1976) Hierarchical organization: a candidate principle for ethology. In *Growing Points in Ethology* (Bateson, P.P. G. and Hinde, R.A., eds), pp. 7–54, Cambridge University Press
108. Seyfarth, R.M. and Cheney, D.L. (2014) The evolution of language from social cognition. *Curr. Opin. Neurobiol.* 28, 5–9
109. Doupe, A.J. and Kuhl, P.K. (1999) Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci.* 22, 567–631
110. Petkov, C.I. and Jarvis, E. (2012) Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4, 12
111. Lilly, J.C. (1965) Vocal mimicry in tursiops: ability to match numbers and durations of human vocal bursts. *Science* 147, 300–301
112. Janik, V.M. (2009) Acoustic communication in dolphins. *Adv. Study Behav.* 40, 123–157

113. Connor, R.C. (2007) Dolphin social intelligence: complex alliance relationships in bottlenose dolphins and a consideration of selective environments for extreme brain size evolution in mammals. *Philos. Trans. R. Soc. B* 362, 587–602
114. Herman, L.M. *et al.* (1993) Responses to anomalous gestural sequences by a language-trained dolphin: evidence for processing of semantic relations and syntactic information. *J. Exp. Psychol. Gen.* 122, 184
115. Marler, P. *et al.* (1986) Vocal communication in the domestic chicken. I. Does a sender communicate information about the quality of a food referent to a receiver? *Anim. Behav.* 34, 188–193
116. Jarvis, E.D. (2006) Selection for and against vocal learning in birds and mammals. *Ornithol. Sci.* 5, 5–14
117. Arnold, K. and Zuberbühler, K. (2006) Language evolution: semantic combinations in primate calls. *Nature* 441, 303–303
118. Căsar, C. *et al.* (2012) The alarm call system of wild black-fronted titi monkeys, *Callicebus nigrifrons*. *Behav. Ecol. Sociobiol.* 66, 653–667
119. Căsar, C. *et al.* (2013) Titi monkey call sequences vary with predator location and type. *Biol. Lett.* 9, 20130535
120. Cheney, D.L. *et al.* (1995) The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning? *J. Comp. Psychol.* 109, 134
121. Rey, A. *et al.* (2012) Centre-embedded structures are a by-product of associative learning and working memory constraints: evidence from baboons (*Papio papio*). *Cognition* 123, 180–184
122. Lew-Williams, C. and Saffran, J.R. (2012) All words are not created equal: expectations about word length guide infant statistical learning. *Cognition* 122, 241–246
123. Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204
124. Fiser, J. *et al.* (2010) Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* 14, 119–130
125. Testolin, A. and Zorzi, M. (2016) Probabilistic models and generative neural networks: towards a unified framework for modeling normal and impaired neurocognitive functions. *Front. Comput. Neurosci.* 10, 73