RECENTS ADVANCES IN NEUROBIOLOGICAL DOMAINS: STATISTICAL LEARNING AND PREDICTIVE PROCESSING FOR LANGUAGE

Mirela Cunha Cardoso Ramacciotti

https://doi.org/10.1590/SciELOPreprints.14880

Submitted on: 2026-01-17
Posted on: 2026-01-27 (version 1)
(YYYY-MM-DD)

AVANÇOS RECENTES EM DOMÍNIOS NEUROBIOLÓGICOS: APRENDIZAGEM ESTATÍSTICA E PROCESSAMENTO PREDITIVO DA LINGUAGEM

RECENTS ADVANCES IN NEUROBIOLOGICAL DOMAINS: STATISTICAL LEARNING AND PREDICTIVE PROCESSING FOR LANGUAGE

Revisão de Literatura

Mirela C. C. Ramacciotti, Doutora, Universidade de São Paulo, Departamento de Psicologia Experimental, São Paulo, mirela-ramacciotti@usp.br, ORCID: https://orcid.org/0000-0002-3109-8376

RESUMO: Pesquisas em neurobiologia da linguagem na primeira infância são fundamentais para entender redes linguísticas maduras. Graças a técnicas recentes de neuroimagem não invasiva, como fNIRS, temos um conhecimento aprimorado sobre o cérebro em desenvolvimento que permite identificar marcadores precoces de resultados linguísticos. Esta revisão narrativa visa analisar diferentes domínios para revelar experimentos e técnicas que abrem caminho para o processamento preditivo no desenvolvimento da linguagem. No domínio da aprendizagem estatística, precisamos especificar padrões de mudança que permitam extrair regularidades do ambiente no desenvolvimento inicial, cientes de que o armazenamento de longo prazo da aprendizagem estatística nas crianças ainda é pouco compreendido. Em relação às variáveis contextuais, há pesquisas recentes que trazem insights sobre como a previsibilidade do comportamento do cuidador estrutura o neurodesenvolvimento de processos específicos de aprendizagem. E sobre restrições de ordem maturacional e de desenvolvimento, exploramos descobertas recentes que expõe dúvidas, como sobre crianças e adultos possuírem habilidades estatísticas linguísticas semelhantes. Portanto, domínios, variáveis e restrições são cruciais para integrar a pesquisa com aspectos fundamentais de como a segregação tende a se desenvolver para integração em domínios cognitivos complexos, como memória e linguagem. Importante ressaltar que o neurodesenvolvimento da aprendizagem, assim como das habilidades linguísticas, continua dependente das experiências iniciais durante janelas de desenvolvimento cruciais. Isso valida uma exploração mais profunda dos processos neurobiológicos que sustentam o desenvolvimento. No final, esperamos oferecer uma avaliação mais detalhada dos avanços recentes sobre como os conceitos de aprendizagem estatística e processamento preditivo influenciam os resultados linguísticos.

PALAVRAS-CHAVE: Aprendizagem Estatística; Processamento Preditivo; Neurodesenvolvimento; Maturação; Habilidades linguísticas

ABSTRACT: Research in language neurobiology in early childhood is key to understanding mature language networks. Thanks to recent non-invasive neuroimaging techniques, such as fNIRS, we have an upgraded knowledge about the developing brain that allows for the identification of early markers of language outcomes. This narrative review aims to look at different domains to unveil experiments and techniques that pave the way for predictive processing in language development. In the domain of statistical learning, we need to specify patterns of change that allow for the ability to extract environmental regularities in early development, while being aware that the long-term storage of statistical learning in children remains poorly understood. In relation to contextual variables, there is recent research that bears insights into how predictability of caregiver behavior scaffolds the neurodevelopment of specific learning processes. And regarding maturational and developmental constraints, we explore recent findings that expose dissent as to whether children and adults have similar linguistic statistical learning abilities. Henceforth, domains, variables, and constraints are crucial to integrate research with foundational aspects of how segregation tends to develop into integration in high cognitive domains, like memory and language. Of note, the neurodevelopment of core learning, like linguistic abilities, relies on infants' experiences during key developmental windows. This validates deeper exploration of neurobiological processes that undergird development. In the end, we hope to enable a more thorough appraisal of recent advances on how statistical learning and prediction frameworks influence language outcomes.

KEYWORDS: Statistical Learning; Predictive Processing; Neurodevelopment; Maturation; Linguistic Abilities

RESUMO PARA NÃO ESPECIALISTAS: Conhecer sobre como o cérebro em desenvolvimento opera nos permite identificar marcadores precoces de resultados linguísticos. Isso pode ser feito com várias técnicas que deixam mais claro processos neurais, como a técnica de neuroimagem chamada fNIRS. Aqui narramos alguns estudos, que usam essas técnicas e se baseiam no domínio da aprendizagem estatística e da ideia de predição, para mostrar evidências de como bebês adquirem a linguagem. Importa nesse contexto saber onde há variação e onde há restrição. Por isso, trazemos estudos sobre interações entre cuidador e bebê e sobre limites definidos pela idade e pelo estágio de desenvolvimento dos bebês ao adquirirem linguagem. Ressaltamos aqui o papel da experiência na identificação e aquisição de regularidades que guiam a aprendizagem implícita. Há períodos cruciais para essa relação que podem influenciar, positiva ou negativamente, os resultados esperados.

Introduction

There is no doubt today that understanding language neurobiology in early childhood is key for how mature language networks operate. Language is a domain of abstract cognitive function (Loo; Pavlick; Feinman, 2026). As such, it engenders neural substrates and mechanisms that lie at the basis of how we think and interact. Thus, it becomes crucial for human development (Scheinost *et al*., 2021). For example, soundwaves are the input that set patterns for speech production – a necessary staple for interaction. Henceforth, neurobiological precursors that detect and segregate speech streams - already present at birth (Kujala *et al*., 2023) and that can be turned into speech production via implicit statistical learning (Wang; Lu; Wu, 2025) - need to be better understood. Those that cater for how language acquisition operates in early years, like caregivers and teachers, should target such information to avoid obstacles to interaction that may emerge across development (Ramacciotti; Bailer; Noro, 2023).

Thanks to recent non-invasive neuroimaging techniques, as fNIRS, we have an upgraded knowledge about the developing brain that allows for the identification of early markers of language outcomes (Gervain, 2014; Olson *et al*., 2025). Importantly, techniques that are user-friendly to children in specific allow for understanding the development of language – a unique human ability - before behavior develops, as comprehension is better understood before production emerges (Olson *et al.,* 2025). While techniques like EEG are precious for time-locked analysis of acoustic and language real-time processing; wearable, non-intrusive devices like fNIRS allow for special precision in analyzing how specific abilities unfold in the brain (Ramacciotti *et al*., 2024). For example, left-dominance in hemispheric engagement for right-handed individuals in language has received much attention. In using FNIRS with a 12 full-term newborns, Pena *et al*. (2013) tested whether there was hemispheric superiority at birth by exposing participants to normal and backward speech or silence while employing fNIRS to gauge brain activation. Results demonstrate that, right after birth, there is sensitivity to linguistic input. Further, it specifically engages the left hemisphere.

Neuroimaging techniques are potent in specifying neuromarkers to evaluate typical and atypical brain development. Indeed, a groundbreaking study (Scheinost *et al*., 2021) used fMRI in a resting state. They performed an analysis of functional connectivity between Broca's and Wernicke's areas – primers of the language network - interhemispherically. Their participants ranged from 30 weeks of gestation through 30 months old. They found fast-developing, strong interhemispheric functional connectivity from the third trimester of gestation through the first postnatal month while intrahemispheric connectivity remained weak till 30 months of age. Promising research may come from studies that use *hyperscanning*, that is, when two different techniques, like fNIRS and EEG, are used together. This review aims at bringing selected examples of studies that employ these techniques, singly or jointly, to shed light in mechanisms for language acquisition.

On par with the importance of language and recent findings of neurobiological substrates, this narrative review analyses different **domains**, like statistical learning (Aslin, 2017) and **concepts**, like prediction (Bubic; von Cramon; Schubotz, 2010), to unveil experiments and findings that have charted predictive processing in language development. Being sensitive to statistical properties of the input guides learning not only in language but across other domains, like perception and emotion (Hackel *et al*., 2016). Therefore, to move forward, we need to perform and acknowledge cutting-edge research. Specifically, in the domain of statistical learning, we need to further specify patterns of change that allow for the ability to extract environmental regularities in early development while aware that the long-term storage of statistical learning in children remains poorly understood. Indeed, although some have found evidence that learning linguistic stimuli did not change across ages (Ravi; Arnon, 2018), much theoretical work has advanced the notion of a linguistic statistical learning advantage in the early years (Moreau *et al*., 2022). The basis for such idea rests on "linguistic entrenchment", i.e., implicit learning of a novel artificial language would suffer from an adult's knowledge of their native language. Henceforth, children, still experimenting with their native language, would be more nimble to contrasts and faster at adaptations required for new representations. If that were the case, then statistical learning in language would differ from other domains (Moreau *et al.,* 2022).

Context shapes development. For language, input drives forward predictive processing necessary for learning. However, across development, brain processing is driven by specialization rather than generalization. In relation to contextual **variables**, there is recent research (Forest *et al*., 2025; Nguyen; Zimmer; Hoehl, 2023) that bears insights into how predictability of caregiver behavior scaffolds the neurodevelopment of specific learning processes. Importantly, as predictability influences statistical learning, would changes in modality, like those between auditory and vocalizations, affect this influence? And regarding maturational **constraints**, would we observe differences between children and adults regarding linguistic statistical learning abilities?

Overall, these are some questions that illustrate how domains, concepts, variables and constraints are crucial to integrate research with foundational aspects of how segregation tends to develop into integration in high cognitive domains, like memory and language. Of note, the neurodevelopment of core learning, like linguistic abilities, remains dependent on infants' experiences during key developmental windows. This validates a deeper exploration of neurobiological processes that undergird development, such as the processes that we have selected for this review. In the end, we hope to enable a more thorough appraisal of recent advances that further how statistical learning and prediction frameworks influence language outcomes from a neurobiologically informed perspective

1.Statistical learning domain

Statistical learning is regarded as a cognitive, domain-general mechanism that enables individuals to extract patterns, of a visual or auditory nature, present in the environment (Aslin; Newport, 2012; Fló *et al.,* 2025; Wang; Lu; Wu, 2025). Nevertheless, it has only very recently been evidenced as a mechanism (Fló *et al.*,2025) despite the idea being not new to the area of neuroscience nor to language development (Saffran, 2009; Saffran *et al.*, 1999).

Empirical evidence that implicit, statistical learning subserves language acquisition across life came from Conway *et al.* (2010). Based on simple recurrent network theories, they hypothesized that implicit learning abilities encode the word order regularities of language and can be used to improve speech perception. To test their hypothesis, they devised experiments that reexamined and extended assumptions at each step. Thus, they performed three experiments: one to test whether implicit learning is associated with long-term knowledge of word predictability in sentences; another to test whether implicit learning is directly associated with knowledge of word order predictability; and a third to confirm the association between implicit learning and knowledge of word predictability. Overall, they showed that sensitivity to sequential structure is significantly correlated with the ability to use knowledge of word predictability. That correlation underlies speech perception even when listening conditions are not ideal. The study, performed with young adults, underscores individual variability on acquiring long-term knowledge of word predictability with an impact on speech perception. Of note, the study sedimented the knowledge that implicit learning is very specific regarding language processing, i.e., we learn statistically to acquire knowledge about the predictability of items in a sequence. More recent research (Schneider *et al.*, 2024), also with an adult population, pinpointed the neural correlates of speech regularities in linguistic statistical learning in the left posterior temporal gyrus. Such examples illustrate that the mechanism of statistical learning subserves language processing across ages.

Saffran (2009) prefers to refer to statistical learning as a model-system that lays two fundamental claims: (i) there are relevant environmental structures that can be detected statistically or in patterns, and (ii) there are organisms that can extract such patterns. Of note, Saffran (2009) underscores that the matching of a pattern that is informative with an individual's mechanism that recognizes and capitalizes on it is not accidental.

Box 1: Stochastic Processes and Language

It has not been too long since the field of language acquisition lay the grounds for an understanding of language as a set of complex and variable patterns that relate in a statistical (or distributional) way (Miller & Selfridge, 1950). Central to this idea - presently moving forward by LLMs which work in tandem with models that implement Bayesian probabilistic language of thought (Loo; Pavlick; Feinman, 2026) - were the experiments performed by Markov (1907/1971) which subserved the groundbreaking works of Shannon (1948). Alexander A. Markov analyzed the distribution of vowels and consonants in Alexander Pushkin's *Eugene Onegin*. In his work, he treated letters as abstract categories thus removing their semantic content. Shannon (1948) used a Markov chain – which is a probabilistic process where the state change relies only on the current state - to propose a statistical model of the sequences of letters in a text. In other words, the process models a number of distinct states and the probabilities that a system would transition from one state to another. To that end, he used a Markov model, which is a mathematical structure, to simulate the statistical structure of a piece of text. What the Markov model provides is a way to predict each letter of the alphabet with an assigned probability (for a practical example, see Sedgewick; Wayne, 2004). Thus, Markov and Shannon turned overt a process that was, until then, covert.

In truth, the infant brain at birth is ready to perform statistical learning that includes (i) detecting and learning patterns from the input, and (ii) generalizing such patterns and applying them to new instances (Aslin; Newport, 2012; Kujala *et al.*, 2023). These are the abilities that enable infants to learn when instruction is not present. Statistical learning is an implicit mechanism (Conway, 2020). Thus, infants learn to categorize sounds and detect word boundaries, recognize words, and separate speech streams without the need for instruction (Kujala *et al*., 2023). By detecting and applying patterns to the novel linguistic input they receive, infants move fast in their language acquisition process within their first year of life.

Infants learn what is predictable or regular in their surroundings. Although the learning is implicit, the aim is clear: to predict changes (Wang; Lu; Wu, 2025). And this is relevant because, by doing so, infants can rely on patterns and direct precious cognitive resources to what surprises them. Notably, throughout life, individuals achieve that by unconsciously processing sensorial information. The noticeable feature in this sort of processing is that it enables us to distinguish between probabilities and regularities. For example, Saffran, Aslin and Newport (1996) evidenced that infants at 8 months of age learned to extract words from a monotone speech stream, i.e., without acoustic boundaries, after a 2-minute-long exposure. Results evidenced statistical learning in using the transition probabilities between syllables. This study was followed by another one, performed by the same group of researchers (Aslin; Saffran; Newport, 1998), that went further in documenting conditional probability in word segmentation. This time round, infants were exposed to trisyllabic nonsense words for 3 minutes to test if they could discriminate between words and part-words. Results indicate that, at 8 months of age, babies rely on transitional probabilities of syllable pairs to segment continuous speech.

Perhaps it did not come as a surprise when studies evidenced statistical learning present in language acquisition before birth (Kujala *et al*., 2023; Wang; Lu; Wu, 2025). Babies in utero are able to detect the melodic cadence of their mothers' language denoting a 'prenatal prosodic bootstrapping' that subserves acquisition (Gervain, 2018). Before birth, sensitivity to mother's voice and speech undergo neural encoding (Kisilevsky *et al*., 2009). The auditory apparatus that subserves this sensitivity needs reckoning.

Webb *et al*. (2015) performed a study that showed experience-dependent plasticity in the auditory cortex. They did that by examining 40 infants born prematurely (between weeks 25 and 32 in utero) during their first month of life. Their experiment consisted of exposing the infants to audio recordings of maternal sounds (voice and heartbeat) versus exposure to routine hospital environmental noise (control condition). Cranial Ultrasound measures were used. Their results showed that infants exposed to maternal sounds had a larger auditory cortex bilaterally compared with controls. Further, they found that bilateral auditory cortex thickness correlated with gestational age. Their study evidenced infants' brain experience-dependent plasticity in the womb.

Notably, precursors to infant's abilities to detect phonemes, words and phrases, and segregate speech streams (from other sounds) underlie their evolving abilities to parse speech units and construct speech hierarchy all the while dealing with the variability present in natural speech (Kujala *et al.*, 2023; Ramacciotti; Bailer; Noro, 2023). For example, sensitivity to rhythm is a precursor to speech segmentation and morphosyntactic acquisition (Gasparini *et al.*, 2021; Goswami, 2022).

2. The concept of prediction

Regarding language processing, incoming sounds are the input that drive forward predictive processing (see Box 2) in the infant brain. When infants detect sound regularities, they encode such patterns in internal models that assign probabilities to stimuli they may encounter. That is how they generate models based on regularities and are able to make inferences about what constitutes speech. For example, a study with newborns (Háden *et al.*, 2015) wanted to investigate predictive processing in their auditory cortex. Based on the premises of predictive coding (see Box 2), researchers used EEG to record neural responses to pitch tones. Their assumption was that if trains were processed in a predictive manner, deviant pitch would generate a mismatch response, that is, a prediction error (see Box 2). More specifically, they tested whether newborns' brains were sensitive to the probability of ascending vs. descending pitch. Their results suggest that, at birth, infants already encode pitch trends within sound sequences predictively. Importantly, their study evidenced that infants display adult-like capabilities when processing non-repetitive sound patterns.

The question of predictive coding subserving language acquisition was tackled in a groundbreaking fNIRS study (Emberson; Richards; Aslin, 2015). They asked whether expectation-based feedback is already present in the sensory cortex of the infant brain. They discovered that the basic neural architecture for sensory feedback is already in place in infancy. Specifically, they evidenced that the infant occipital cortex responds to a prediction error between visual expectation and current sensory input.

Another study that aimed at the role of experience in predictive processing regarding acoustic representations also used fNIRS (Emberson *et al.*, 2019). Their aim was two-fold: (i) to gauge infant brain's sensitivity to statistical learning, and (ii) to check if the neural responses to repetition and variability vary as a function of their probabilistic nature. Their study used auditory repetition suppression to probe whether the infant brain would be sensitive to top-down mechanisms. They employed two experimental conditions conditioned on variability (variable and uniform blocks) with 6-month-old infants. Results showed that when sensory input matches expectations, the infant brain exhibits an attenuated response. To sum, they evidenced that prediction can modulate responses to sensory input at 6 months of age.

Brains resort to specialization to perform actions and better support adaptive behaviors (see Box 2). This is not different in language processing. Predictive processing as a principled framework that enables perception and cognition makes viable and actionable findings that open up our understanding of specific language processing. Case in point, the more extensive a child's repertoire, the better they get at learning. Once studies on predictive processing show how word recognition and acquisition get faster with prediction, the more options the environment may offer to boost children's inferencing process. For example, Ylinen *et al.* (2017) examined EEG responses to syllables from infants at 12 and 24 months of age. Their goal was to test prediction of word-expectancy. They hypothesized that familiarity with a preceding word context would generate more efficiency in predicting the word ending. They found evidence for predictive processing at both ages and of prediction errors for unfamiliar

Box 2: Predictive Processing

Our brains are attuned to our survival. To that end, they use up energy to guarantee our adaptation. Thus, being able to predict future scenarios becomes a necessity. Predictive Processing is a pervasive designation of a probabilistic processing performed by the brain. In this framework, neural networks carry out an inferencing process based on a distribution of probabilities. Resting on a Bayesian approach, the brain weighs probabilities to make inferences about the causes of certain phenomena. Predictions are probabilities generated *a posteriori* to infer the possible causes of new input or information. The process accounts for the constant updating of one's internal model of their environment (generative model). Neural responses encode what is predictable and what is not. This mismatch accounts for prediction errors that update the internal model. By making top-down predictions and using bottom-up prediction errors, the overall aim of predictive processing is to make sensory data less surprising, thus minimizing the amount of energy spent, i.e., of free energy. Surprise accounts for the degree of improbability of an input. Entropy is the average of surprise overtime denoting how much energy is dispersed (Friston; Kiebel, 2012; Ramacciotti, 2026).

words that correlated with younger infants' vocabulary scores. In sum, predictive coding seems to work as an accelerator of word learning.

3. Contextual variables: caregiver behavior

Development is not linear nor equally achieved. This means that there are contextual influences that may upend universal communication structures, like turn-taking, happening very early on between caregiver and infant. In this section, we review two relevant studies that delved upon such structure to revel intriguing findings.

Nguyen, Zimmer and Hoehl (2023) investigated turn-taking in dyads of caregivers and infants at 4-6 months old with a neuroimaging hyperscanning design. They wanted to understand the mechanism of early turn-taking and evaluate possible links between that structure and developmental outcomes. Specifically to the objectives of the present review, they brought findings on interhemispheric connectivity and vocabulary size. As predicted, their results showed that, especially early on in the interaction, more turn-taking yielded more neural synchrony – a neural signature for social cognition. Interestingly, the higher the number of infants' turns, the higher the interhemispheric connectivity in medial prefrontal regions. This means that more frequent interactions between mother and baby matters for brain maturation. In relation to vocabulary size, total number of turns between mother and baby when they were 4 to 6 months old , especially infants' turns, predicted a larger vocabulary size at 2 years of age.

Resting on the premise of greater brain malleability during infancy, Forest *et al*. (2025) proposed a longitudinal experimental study with infants to test the hypotheses that early linguistic experiences with primary caregivers shape the operation of core learning processes. They investigated whether a more predictable early environmental input develops infant's learning predictability. Their means to investigate this idea was to record interactions between primary caregivers and their infant. The study was carried out with a South African population of babies between 2 and 6 months of age and their caregivers. To test, they recorded dyadic interactions between primary caregivers and infants (n= 103, 46 females, 57 males) in a first visit. Their metrics followed caregiver vocalizations while their measure of unpredictability was caregiver entropic behavior. Entropy (see Box 2) was measured via annotated caregiver behavior in recordings of naturalistic interactions between caregiver and infant. The hand-made annotations, carried out separately, were based on a protocol (Davis et al., 2017) that captured sensory-based behaviors (vocalizing as a measure of auditory signal, holding and touching the baby for tactile signals, holding and pointing to an object in the room as visual signals).

Additionally, they annotated attentional behavior, via eye-gazing, included as a covariate. Their Cohen's Kappa of 0.75 was extracted to demonstrate high inter-rater reliability.

To model the entropy of caregiver behavior, Forest *et al.* (2025) used Markov models (see Box 1) that accounted for transitions between different behavioral states and enabled quantifications of consistency (or predictability) of transitions. After applying a formula (Shannon Conditional Entropy) that models uncertainty based on overall frequency and in-between transitions (a Bayesian approach), researchers were able to obtain the "conditional entropy". This denotes the unpredictability of a caregiver's future behavior based on their current state, i.e., the higher the result, the lower the predictability of that behavior. Researchers included two entropy measures: for overall entropy of caregiver behavior and conditioned on caregiver vocalizing. Their aim was to capture transitions between behavioral states.

After five months, during a second visit, Forest et al. (2025) tested the same participants with an auditory statistical learning task and took EEG measures. EEG recordings were specifically used to detect differences in predictability regarding tones via their neural responses. In the described study, EEG data represented infants' measure of learning of a tone stream structure. Their results were based on a comparison of neural responses obtained from auditory information presented during the task with predictable and unpredictable items.

EEG responses are registered in event-related potentials (ERP), a technique that allows the detection of electrical brain responses time-locked to events of a sensory, cognitive, or motor nature. Such potentials, depicted as waveforms, are registered in peaks and troughs as negative (N) or positive (P) respectively. Their voltage is plotted in reverse, that is, with negative waves going up (Woodman, 2010). In Forest *et al.*'s study, a difference in latency of N1 and amplitude of P2 brain oscillations was analyzed following previous results that detected a difference denoting predictability (Pierce; Camody; Nelson III, 2021). N1 was targeted because it translated the detection and discrimination of auditory stimuli while P2 registered learning (Forest *et al.*, 2025). By targeting both responses, the study afforded a verifiable way to test predictability of early linguistic input based on statistical learning.

Overall, Forest *et al.*'s study could answer their hypotheses, of early linguistic experience with high predictable input shaping future learning, by targeting a central relationship; that between entropy (from visit one) and predictability (from visit two). To that end, they secured that enough variability (for entropy) and sensitivity to statistical structure (for predictability) were present. Interestingly, they were also able to (i) detect whether attention, via eye gazing, would influence the targeted relationship and (ii) discriminate between amount of exposure versus structural predictability as an influence on statistical learning by rating frequency of caregiver vocalizations.

Forest *et al.* (2025) found that entropy among caregivers varies greatly and provided evidence of a change in learning. Of note, they found that babies' statistical learning is related to caregiver entropy. Caregivers that were highly predictable (displayed low entropy) in their vocalizations during the naturalistic interactions with their infants accounted for greater, positive learning changes. Also, the more the infant gazed at their caregiver, the better the learning with a detected effect of caregiver vocalizations on auditory learning. In sum, sensory modality is relevant to how predictability influences statistical learning.

4. Maturational and developmental constraints

It has been acknowledged, for quite some time, that every language has phonotactic constraints. This means that there are particular restrictions on which sequences of sounds are permissible and which are not in the words of a certain language. To illustrate, three studies developed in the early 90s are described followed by more recent findings.

In 1992, Kuhl *et al*. tested a group of American and Swedish infants. They were exposed to two sets of vowel stimuli that contained the American English /i/ and Swedish /y/. To make their case clear, researchers included an ideal instance of the vowel – termed the prototype meant to serve as a perceptual magnet – and added 32 variants separated so as to form four rings of eight stimuli each around the prototype. They found that 6-month-olds showed an altered phonetic perception. The reason pointed to exposure to the regularities of their specific language. In 1993, Jusczyk *et al.* illustrated how English differs from Dutch in not allowing words to begin with consonant sequences such as /zw/ or /vl/. What those researchers underscored then was that at 9 months, babies would favor only words that began with permissible sound sequences. In 1995, Jusczyk and Aslin demonstrated that at 7.5 months of age infants did listen longer to familiarized word targets as opposed to unfamiliarized ones. In tandem, these early findings, among others, paved the field's perception of temporal constraints on infant's sensibility. Also, they highlighted a maturational constraint; age matters to how statistical learning evolves over time and allows infants to detect distributional regularities within sequential streams as attested by Wang, Lu, and Wu (2025). More recent research places phonotactic learning at around 5 months of age (Sundara; Breiss, 2020) while at 9 months, infants already access low-frequency phonotactic combinations (Archer; Czarnecki; Curtin, 2021).

Prior research established that expectations about input modulate statistical learning for adults (Dal ben; Souza; Hay, 2021). However, there remained in the field questions whether children processed new input –statistically pre-defined – using their knowledge about the probability of syllable transitions, much like their older counterparts do. Stärk, Kidd and Frost (2022) asked that question in a study with a sample of German children (aged between 7 and 9 years). Their findings confirmed their hypothesis that children who had acquired the patterns in their natural language used them to process new linguistic input, much like adults do. Additionally, they confirmed prior research showing that children are fast at acquiring new words when exposed to input regularities. Interestingly, upon observing that children's recall across tasks remained stable, they concluded that there may indeed be a critical period for language learning, when children learn faster and better in comparison to adults (Newport, 1990).

Also in relation to age, there is a constraint regarding rhythm. From birth, it has been long acknowledged that infants can discriminate between languages by time-locked classes, i.e., they differentiate stress from syllable and mora rhythms (Mehler *et al*., 1988). A meta-analysis carried out to quantify language discrimination skill changes over the first year of life found that age is crucial for preference of non-native languages similar in rhythm to infants' native language (Gasparini *et al*., 2021).

Differences in the abilities that allow for semantic and phonological processing in children and adults have engaged recent research efforts. Developmental constraints under the predictive processing framework inspired a recent study (Angulo-Chavira; Castellón-Flores; Arias-Trejo, 2025). Researchers examined the emerging predictions in meaning and sound with 98 Mexican infants equally distributed by age groups (18, 24, and 30 months old). Their testing paradigm was based on preferential looking task spanning three experiments. The task consisted of constrained and neutral sentences paired with two images (a target and a distractor). In the first experiment, researchers tested word prediction accuracy using predictable sentences (e.g., "The hen lays eggs") and unpredictable sentences (e.g., "My mom bought eggs") paired with two images: a target image (e.g., eggs) and a distractor. In the second experiment 2, they tested semantic prediction with a semantic competitor (e.g., egg-jelly). In the third experiment, they tested phonological prediction with a phonological competitor (e.g., [egg]-[bone]) . Results showed developmental constraints. Specifically, prediction in preferential looking emerged at 18 months. By 24 months, prediction for semantic cues emerged. At 30 months of age, infants showed phonological prediction. The study suggests a developmental pattern in the emergence of relationships; semantic predictions would lay the basis for phonological predictions.

Another study used EEG to search for differences in statistical learning between children and adults (Moureau *et al*., 2022). They found that developmental changes not related to language per se, like those in the attentional domain or auditory habituation, may explain differences in neural entrainment at word- and syllable-levels.

Figure 1: Timeline of relevant findings related in this review for language acquisition regarding statistical learning and predictive processing.

Specifically, adults' neural response at the word level relative to syllable level increases in relation to children's. Conversely, at the syllable level, children's neural entrainment remains stable while adults' decrease. Regardless, the convergence of their findings for neural entrainment at word and syllable levels show robustness across ages with noticeable increases at word-level due to language exposure. Differences, then, seem indeed attributable to development. Of note, their groundbreaking research counteracts the notion that statistical learning improves throughout childhood. They conclude that, given the robustness of statistical learning across the life course, children's advantage for language learning may not be attributed to a linguistic advantage in this domain.

5. Conclusions

Learning regularities that subserve abstract domains, like language, seem to rely on predictive processing. Based on patterns, human brains still in in utero extract probabilities via unconscious, implicit mechanisms. Statistical learning underlies tone, speech and language acquisition processes that develop rapidly in the early years. Of note, individual variability underlays the neuromarkers attesting the emergence of distributional mechanisms in the succession of maturational and developmental stages and constraints (see Figure 1).

Making use of a largely researched framework, that of Predictive Processing, seems relevant for present and future research endeavors in neurobiologically-informed language acquisition. Literature is rife with examples of how predictability enhances efficiency. This is important because it offloads cognition, allowing for a more efficient language acquisition. From a very early age, infants employ top-down mechanisms that seeks adaptation to their experience in developing language.

In tandem with a need to refine understanding of prediction in language processing, developmental constraints lay bare the dynamic nature of prediction. As infants get exposed to and experiment with language, top-down predictions are used to refine bottom-up representations of much-needed linguistic input to update generative models of language. Importantly, the exponential increase in linguistic abilities that are observable in typical developmental trajectories (Figure 1) seem to be linked to the amount and quality of exposure, especially to caregiver vocalizations. That confers a heightened attention to individual characteristics. As noticed by Stärk, Kidd and Frost (2022), children differ in their ability to learn statistically according to their language proficiency. Thus, experience with linguistic input matters to how infants nurture the neurobiologically mechanism of statistical learning that subserves language acquisition.

References

ANGULO-CHAVIRA, Armando Quetzalcóatl; CASTELLÓN-FLORES, Alejandra Mitzi; ARIAS-TREJO, Natalia. Hierarchical prediction in toddlers: Semantic and phonological development. **Journal of Memory and Language**, v. 145, p. 104688, 2025. https://doi.org/10.1016/j.jml.2025.104688

ARCHER, Stephanie L.; CZARNECKI, Natalia; CURTIN, Suzanne. Boosting the input: 9-month-olds' sensitivity to low-frequency phonotactic patterns in novel wordforms. **Infancy**, v. 26, n. 5, p. 745-755, 2021. https://doi.org/10.1111/infa.12423Digital Object Identifier (DOI)

ASLIN, Richard N. Statistical learning: A powerful mechanism that operates by mere exposure. **Wiley Interdisciplinary Reviews: Cognitive Science**, v. 8, n. 1-2, p. e1373, 2017. https://doi.org/10.1002/wcs.1373

ASLIN, Richard N.; NEWPORT, Elissa L. Statistical learning: From acquiring specific items to forming general rules. **Current directions in psychological science**, v. 21, n. 3, p. 170-176, 2012. https://doi.org/10.1177/0963721412436806

ASLIN, Richard N.; SAFFRAN, Jenny R.; NEWPORT, Elissa L. Computation of conditional probability statistics by 8-month-old infants. **Psychological science**, v. 9, n. 4, p. 321-324, 1998. https://doi.org/10.1111/1467-9280.00063

BUBIC, Andreja; VON CRAMON, D. Yves; SCHUBOTZ, Ricarda I. Prediction, cognition and the brain. **Frontiers in human neuroscience**, v. 4, p. 1094, 2010. https://doi.org/10.3389/fnhum.2010.00025

CONWAY, Christopher M. How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. **Neuroscience & Biobehavioral Reviews**, v. 112, p. 279-299, 2020. https://doi.org/10.1016/j.neubiorev.2020.01.032

CONWAY, Christopher M. et al. Implicit statistical learning in language processing: Word predictability is the key. **Cognition**, v. 114, n. 3, p. 356-371, 2010. https://doi.org/10.1016/j.cognition.2009.10.009

DAL BEN, Rodrigo; SOUZA, Debora de Hollanda; HAY, Jessica F. When statistics collide: The use of transitional and phonotactic probability cues to word boundaries. **Memory & Cognition**, v. 49, n. 7, p. 1300-1310, 2021. https://doi.org/10.3758/s13421-021-01163-4

DAVIS, Elysia Poggi et al. Exposure to unpredictable maternal sensory signals influences cognitive development across species. **Proceedings of the National Academy of Sciences**, v. 114, n. 39, p. 10390-10395, 2017. https://doi.org/10.1073/pnas.1703444114

EMBERSON, Lauren L. et al. Expectation affects neural repetition suppression in infancy. **Developmental cognitive neuroscience**, v. 37, p. 100597, 2019. https://doi.org/10.1016/j.dcn.2018.11.001.

EMBERSON, Lauren L.; RICHARDS, John E.; ASLIN, Richard N. Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. **Proceedings of the National Academy of Sciences**, v. 112, n. 31, p. 9585-9590, 2015. https://doi.org/10.1073/pnas.1510343112

FLÓ, Ana et al. Statistical learning beyond words in human neonates. **Elife**, v. 13, p. RP101802, 2025. https://doi.org/10.7554/eLife.101802

FOREST, Tess Allegra et al. Early caregiver predictability shapes neural indices of statistical learning later in infancy. **Developmental Science**, v. 28, n. 1, p. e13570, 2025. https://doi.org/10.1111/desc.13570

FRISTON, Karl; KIEBEL, Stefan. Predictive coding under the free-energy principle. **Philosophical transactions of the Royal Society B: Biological sciences**, v. 364, n. 1521, p. 1211-1221, 2009. https://doi.org/10.1098/rstb.2008.0300

GASPARINI, Loretta et al. Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis. **Cognition**, v. 213, p. 104757, 2021. https://doi.org/10.1016/j.cognition.2021.104757

GASPARINI, Loretta et al. Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis. **Cognition**, v. 213, p. 104757, 2021. https://doi.org/10.1016/j.cognition.2021.104757

GERVAIN, Judit. The role of prenatal experience in language development. **Current opinion in behavioral sciences**, v. 21, p. 62-67, 2018. https://doi.org/10.1016/j.cobeha.2018.02.004

GOSWAMI, Usha. Language acquisition and speech rhythm patterns: an auditory neuroscience perspective. **Royal Society Open Science**, v. 9, n. 7, p. 211855, 2022. https://doi.org/10.1098/rsos.211855

HACKEL, Leor M. et al. On the neural implausibility of the modular mind: evidence for distributed construction dissolves boundaries between perception, cognition, and emotion. **Behavioral and Brain Sciences**, v. 39, 2016. https://doi.org/10.1017/S0140525X15002770

HÁDEN, Gábor P. et al. Predictive processing of pitch trends in newborn infants. **Brain research**, v. 1626, p. 14-20, 2015. https://doi.org/10.1016/j.brainres.2015.02.048

JUSCZYK, Peter W. How infants begin to extract words from speech. **Trends in cognitive sciences**, v. 3, n. 9, p. 323-328, 1999.Jusczyk, P.W. and Aslin, R.N. (1995) Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1–23. https://doi.org/10.1016/S1364-6613(99)01363-7

JUSCZYK, Peter W. et al. Infants′ sensitivity to the sound patterns of native language words. **Journal of memory and language**, v. 32, n. 3, p. 402-420, 1993.Kisilevsky, B. S., Hains, S. M., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., ... & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, *32*(1), 59-71. https://doi.org/10.1006/jmla.1993.1022

KUHL, Patricia K. et al. Linguistic experience alters phonetic perception in infants by 6 months of age. **Science**, v. 255, n. 5044, p. 606-608, 1992. https://doi.org/10.1126/science.173636

KUJALA, Teija et al. Prerequisites of language acquisition in the newborn brain. **Trends in Neurosciences**, v. 46, n. 9, p. 726-737, 2023. https://doi.org/10.1016/j.tins.2023.05.011

LOO, Alyssa; PAVLICK, Ellie; FEIMAN, Roman. LLMs model how humans induce logically structured rules. **Journal of Memory and Language**, v. 146, p. 104675, 2026. https://doi.org/10.1016/j.jml.2025.104675

MARKOV, Andrey. Extension of the limit theorems of probability theory to a sum of variables connected in a chain. **Dynam Probabilist Syst**, v. 1, p. 552, 1971. Orginal Work from 1907.

MEHLER, Jacques et al. A precursor of language acquisition in young infants. **Cognition**, v. 29, n. 2, p. 143-178, 1988. https://doi.org/10.1016/0010-0277(88)90035-2

MILLER, George A.; SELFRIDGE, Jennifer A. Verbal context and the recall of meaningful material. **The American journal of psychology**, v. 63, n. 2, p. 176-185, 1950. https://doi.org/10.2307/1418920

MOREAU, Christine N. et al. No statistical learning advantage in children over adults: Evidence from behaviour and neural entrainment. **Developmental cognitive neuroscience**, v. 57, p. 101154, 2022.NEWPORT, Elissa L. Maturational constraints on language learning. **Cognitive science**, v. 14, n. 1, p. 11-28, 1990. https://doi.org/10.1016/j.dcn.2022.101154

NGUYEN, Trinh; ZIMMER, Lucie; HOEHL, Stefanie. Your turn, my turn. Neural synchrony in mother–infant proto-conversation. **Philosophical Transactions of the Royal Society B**, v. 378, n. 1875, p. 20210488, 2023. https://doi.org/10.1098/rstb.2021.0488

OLSON, Halie A. et al. Utilizing functional neuroimaging to study early language development. **Developmental cognitive neuroscience**, p. 101641, 2025. https://doi.org/10.1016/j.dcn.2025.101641

PENA, Marcela et al. Sounds and silence: an optical topography study of language recognition at birth. **Proceedings of the National Academy of Sciences**, v. 100, n. 20, p. 11702-11705, 2003. https://doi.org/10.1073/pnas.1934290100

PIERCE, Lara J.; TAGUE, Erin Carmody; NELSON III, Charles A. Maternal stress predicts neural responses during auditory statistical learning in 26-month-old children: an event-related potential study. **Cognition**, v. 213, p. 104600, 2021. https://doi.org/10.1016/j.cognition.2021.104600

RAMACCIOTTI, Mirela CC. Introduction to Emotions: An ideological, historical, and transactional overview. *In* WATSON *et al*. **Embodied Emotion in Artificial Psychology**. Routledge. To be released in April, 2026.

RAMACCIOTTI, Mirela CC; BAILER, Cyntia; NORO, Grazielle. Overview of Language Acquisition for an Efficient Inclusion. **Organon**, v. 38, n. 76, 2023. https://doi.org/10.22456/2238-8915.135031

RAMACCIOTTI, M. C. C. et al. Left OFC activation in fNIRS during an inhibitory control task in an early years sample: integrating stress responses with cognitive function and brain activation. **Developmental Neuroscience**, v. 47, n.2, p. 81-97, 2024. https://doi.org/10.1159/000539023

RAVIV, Limor; ARNON, Inbal. The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. **Developmental science**, v. 21, n. 4, p. e12593, 2018. https://doi.org/10.1111/desc.12593

SAFFRAN, Jenny R. What can statistical learning tell us about infant learning? *In* WOODWARD. A.; NEEDHAM, A. **Learning and the infant mind**. Oxford University Press, 2009, p. 29-46.

SAFFRAN, Jenny R. et al. Statistical learning of tone sequences by human infants and adults. **Cognition**, v. 70, n. 1, p. 27-52, 1999.https://doi.org/10.1016/s0010-0277(98)00075-4

SCHEINOST, Dustin et al. Functional connectivity for the language network in the developing brain: 30 weeks of gestation to 30 months of age. **Cerebral cortex**, v. 32, n. 15, p. 3289-3301, 2022. https://doi.org/10.1093/cercor/bhab415

SEDGEWICK, Robert; WAYNE, Kevin. **Markov Model of Natural Language for COS 126**. 2004. Disponível em https://www.cs.princeton.edu/courses/archive/fall18/cos126/assignments/markov/ . Acesso em: 15 jan. 2025.

SHANNON, Claude E. A mathematical theory of communication. **The Bell system technical journal**, v. 27, n. 3, p. 379-423, 1948. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

STÄRK, Katja; KIDD, Evan; FROST, Rebecca LA. The effect of children's prior knowledge and language abilities on their statistical learning. **Applied Psycholinguistics**, v. 43, n. 5, p. 1045-1071, 2022. https://doi.org/10.1017/S0142716422000273

SUNDARA, M.; BREISS, C. 5-month-olds are sensitive to phonotactic patterns in their native language. In: **45th Annual Boston University Conference on Language Development.** https://labphon. org/labphon17/searchable-programme. 2020.

WEBB, Alexandra R. et al. Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. **Proceedings of the National Academy of Sciences**, v. 112, n. 10, p. 3152-3157, 2015. https://doi.org/10.1073/pnas.1414924112

WANG, Yuyang; LU, Li; WU, Meiyun. Statistical learning across cognitive and affective domains: a multidimensional review. **Frontiers in Integrative Neuroscience**, v. 19, p. 1460471, 2025. https://doi.org/10.3389/fnint.2025.1460471

WOODMAN, Geoffrey F. A brief introduction to the use of event-related potentials in studies of perception and attention. **Attention, Perception, & Psychophysics**, v. 72, n. 8, p. 2031-2046, 2010. https://doi.org/10.3758/APP.72.8.2031

YLINEN, Sari et al. Predictive coding accelerates word recognition and learning in the early stages of language development. **Developmental science**, v. 20, n. 6, p. e12472, 2017. https://doi.org/10.1111/desc.12472

Conflict of Interest Statement: None

Link to *Preprint*

Research Protocol and Pre-Registration: Not applicable

Statement of Data Availability: The data used for the review are contained in this instrument. There was no experimental research for the generation of specific data.

This preprint was submitted under the following conditions:

- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a Creative Commons CC-BY license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.