The impact of generative linguistics on psychology: Language acquisition, a paradigm example

Caterina Marino
Integrative Neuroscience and Cognition Center
(INCC UMR8002); Université Paris Descartes
cateemar@gmail.com

Judit Gervain
Integrative Neuroscience and Cognition Center
(INCC UMR8002); CNRS, Paris
judit.gervain@parisdescartes.fr

Abstract: Noam Chomsky's early work was at the core of the "cognitive revolution" in the 1950s-60s, leading to a paradigm shift from a behavioralist to a mentalistic approach to human psychology. Central to this revolution has been the question of how infants learn language. Here, we provide an overview of how the generative enterprise has shaped research on language acquisition over the last decades. We argue that a large body of empirical knowledge about infants' representation of grammar has accumulated. Many of these questions would most likely not have been investigated empirically without the impetus of such a mentalistic approach.

Keywords: language acquisition; generative linguistics; poverty of the stimulus; logical problem of language acquisition; bootstrapping theories of language acquisition

1. Introduction

Noam Chomsky's early work and generative linguistic theory more generally were at the core of the "cognitive revolution" in the 1950s–60s, leading to a paradigm shift from a behavioralist to a more mentalistic approach to human psychology. Central to cognitive science, a discipline born out of this initial movement is the question of how human infants learn their native language(s). Here, we provide an overview of how Chomsky's work and the generative enterprise have shaped research on language acquisition over the last decades. We argue that under the influence of generative grammar, several surprising and highly relevant findings about infants' early knowledge of grammar have been revealed, many of which would most likely

not even have been investigated empirically without the impetus of such a mentalistic approach. We also highlight that more recently, these advances have led to a new, integrative perspective, going beyond the original theoretical dichotomy of nature (nativism) vs. nurture (empiricism).

2. Language acquisition as a logical problem, innateness as a solution

The problem of accounting for language acquisition appeared very early in Noam Chomsky's work. In his review of Skinner's *Verbal behavior*, Chomsky (1959) draws a parallel between the infants' task of learning a language and the linguist's job of describing the grammar of an unknown language. Both the child and the linguist need to derive a rule system from the finite set of linguistic exemplars they encounter. This, Chomsky, argues is a logical impossibility. As the induction problem, a well-known mathematical phenomenon, suggests, for any finite dataset, there is an infinite number of underlying rules that could generate it. It is thus impossible for the learner to know which rule system to select as the "true" grammar of the target language. For language acquisition, this would mean that even if infants settled on a given grammar, there was no guarantee that any two infants settled on the same grammar and would thus end up learning the same language. This, of course, is contrary to empirical fact, as all typically developing infants acquire the language(s) spoken in their linguistic community.

This "poverty of the stimulus" problem is further exacerbated because the linguistic input infants receive is noisy, containing within- and between-speaker variation, errors, omissions, false starts etc. Even more importantly, infants only rarely receive negative evidence (Marcus 1993), i.e., information that a sentence or a form is not part of the grammar, i.e., is ungrammatical, because adults usually only correct factual or objective mistakes in children's speech (e.g., *This is a tiger* when it is actually a lion), but not grammatical errors (e.g., *Tom want chocolate*) (Braine 1971). Further, infants never receive the kind of explicit teaching about the rules of grammar that second language learners often get.

If language cannot be learned from external input alone due to the infinite number of possible grammars, argues Chomsky, an innate system, a "Language Acquisition Device" needs to exist to narrow down the logical space of possible grammars, thereby helping the infant converge on the appropriate grammar. "The fact that all normal children acquire essentially comparable grammars [...] suggests that human beings are [...] designed to do this, with [a] "hypothesis-formulating" ability of unknown charac-

ter and complexity" (Chomsky 1959, 49). This mentalistic approach went against the behaviorist account proposed for instance in Skinner's *Verbal behavior* (Skinner 1957), which framed language learning as a stimulus-response phenomenon, the most important learning mechanism of which is imitation.

The innate nature of the language faculty implies that it is grounded in our genetic endowment. This approach moves away from considering language a cultural object that varies infinitely across populations (e.g., Boas 1940) and places it in the domain of human biology, universal across our species.

The poverty of stimulus argument is not without its critiques. Pullum and Scholz (2002), for instance, provide empirical arguments suggesting that the infant's input may not be as poor in relevant positive evidence as generative linguists usually claim. One often cited example for poverty is stimulus arguments (Crain & Nakayama 1987; Crain 1991) is the structure-sensitivity of question formation when a complex sentence with an embedded relative clause is present. It has been argued that while infants encounter many examples of question formation for simple main clauses (*You are there → Are you there?*), instances of complex cases involving embedded relatives are rare (*The dog that is in the corner is hungry → Is the dog that is in the corner __ hungry?*, and not **Is the dog that __ in the corner is hungry?*). Generativist accounts have argued that sentences of the type *Is the dog that is in the corner __ hungry?*, constituting positive evidence that question formation involves the main verb (main auxiliary) and not the sequentially first one, are rare in young learners' input, yet children typically for these complex questions correctly from an early age (Crain & Nakayama 1987). By contrast, Pullum and Scholz (2002) show that such sentences do appear in the input, citing examples from literature (e.g., William Blake's "Tiger": *Did He who made the lamb make thee?*), the Wall Street Journal corpus, and most relevantly in the CHILDES corpus. In general, the poverty of stimulus argument has given a strong impetus to empirical studies of the input young learners receive. The CHILDES corpus (MacWhinney 2000), one of the largest annotated written and audio(visual) databases of infant-adult and infant-infant interactions, has to a large extent grown out of efforts to better understand the input. Despite this attention and the growing body of available corpus data, no definitive empirical evidence exists to settle the issue of whether the input provides sufficient (positive) evidence for learning. Further work in this regard is thus needed.

3. Formal theories of grammar and language acquisition

Chomsky has proposed several formal theories of Universal Grammar (UG; e.g., Chomsky 1957; 1965; 1981; 1995), the innate mental system that contains the logical space of all possible grammars and as such underlies all human languages. One formalism, the Principles and Parameters model (P&P; Chomsky 1981) offers a framework in which language acquisition is particularly easy to operationalize. The P&P model argues that UG consists of principles, universal features that characterize all human languages, and parameters, variables often conceptualized as (binary) switches that account for cross-linguistic variation. For instance, it is a principle of UG that each sentence must have a subject. A classical example of a parameter is the *pro*-drop or null subject parameter. In languages, e.g., Italian, Hungarian, in which this parameter takes a positive value, the subject of the sentence can be a phonologically empty, null element (*pro*), e.g., Italian: *Piove* rain.3sg 'It rains', whereas in languages with a negative value, e.g., English, French, no empty subjects are allowed (*Rains).

The P&P model readily accounts for language acquisition. Principles raise no learning problem, as they are universal across all languages, and they are innate. Parameters, which account for the observable variation across languages, are the real targets of language acquisition: the learner's task is set the values of the parameters to the ones that characterize the native language. This is a systematic and readily operationalizable view of language acquisition.

However, it leaves open one question: what are the triggers for parameter setting? In other words, parameters and the linguistic variables and categories that constitute them, such as *pro*, V(erb) etc., are abstract mental entities. How does the infant know what corresponds to them in the concrete, physical input she receives? This question was formulated by Pinker (1984) as the "linking problem".

Bootstrapping theories offer a possible solution. They hypothesize that infants learn about certain abstract structural properties of their native language using perceptually available surface cues that are correlated with the abstract properties (Morgan & Demuth 1996). In English, for instance, disyllabic nouns typically have initial stress, while verbs have final stress, e.g., **ré**cord (N) vs. rec**órd** (V). Thus the stress pattern of a word can help infants determine its lexical category even if the word is unfamiliar and its meaning is unknown. Several types of bootstrapping models have been proposed on the basis of the different surface cues that infants may use, e.g., semantic bootstrapping (Pinker 1984), syntactic bootstrapping

(Gleitman & Landau 1994; Lidz et al. 2003), prosodic bootstrapping (Morgan & Demuth 1996; Nespor 1990; Nespor et al. 2008; Gervain & Werker 2013) and frequency-based bootstrapping (Nespor et al. 2008; Bernard & Gervain 2012). The specific contributions of each bootstrapping model will be discussed in greater detail in section 4.2.4 below.

This nativist account of language acquisition is often contrasted in the literature by empiricist accounts arguing that statistically-based, general purpose or other non-linguistic, e.g., social, mechanisms can account for language acquisition without the need to assume innate, language-specific mental contents. It is impossible to give an exhaustive account of these proposals here. We would, nevertheless, like to mention a few influential and particularly relevant examples.

One tradition argues that there is enough statistical information (e.g., frequency of occurrence and co-occurrence of different linguistic units) in the input to provide robust patterns, which young learners can pick up on to learn grammar. Thus, connectionist (Elman et al. 1996) and more recently machine learning (LeCun et al. 2015) modelers have built networks and algorithms that can learn structural regularities from different linguistic inputs with high accuracy. It has also been shown experimentally that very young infants are sensitive to statistical regularities in their input, such as frequency and conditional probability cues (Saffran et al. 1996 and subsequent work). This statistical learning based account has produced remarkable results. Very young infants have been shown to rely on transitional probabilities for word segmentation and other linguistic tasks under a wide variety of circumstances from the earliest ages (e.g., Teinonen et al. 2009; Pelucchi et al. 2009; Lany & Saffran 2010). Today's deep learning algorithms (LeCun et al. 2015) are able to recognize and label large quantities of images, perform intelligent semantic searches, and are regularly used by our web browsers and smart phones for a large variety of purposes. The accuracy of these algorithms is often very high, and matches human performance. However, very tellingly, their errors reveal non-human-like performance: an image categorization algorithm labeled a sofa with a leopard fabric as a leopard. In a systematic review, Marcus (2018) argues that while deep learning and other machine learning algorithms are very powerful, they differ from human cognition in many ways, suggesting that the underlying computational mechanisms are different. For instance, deep learning algorithms need massive amounts of training data, while infants can learn from few instances (e.g., they can learn a word form and its meaning from just a single instance of exposure, a mechanism known as zap-mapping or fast mapping). Deep learning algorithms are

task-specific and close-ended. A network trained on image recognition will do poorly on phoneme categorization, whereas human intelligence is flexible and allows for generalizations. And most importantly for the current discussion, these algorithms cannot handle hierarchical structure like the one found in human syntax. Marcus (2018), therefore, suggests that the best approach, most similar to human cognition and even more powerful than the current algorithms, would be to augment the current models with symbolic representations allowing them to meet some of the challenges mentioned above.

Another empiricist approach, inspired by usage-based theories of language processing (Bybee & Hopper 2001), argues that statistical information in the input complemented with social learning mechanisms such as joint attention between an infant and a caregiver in a given communicative situation best explain language acquisition (Tomasello 2000). According to this approach, the infant picks up common sequences from the input, with at most one open slot in them (e.g., *Where is the __ ?*) and can initially only insert a limited set of items in this open slot. She then proceeds by gradually generalizing from these initially limited and item-based templates to construct an increasingly abstract grammar. In so doing, she relies to a large extent on her advanced social abilities, which allow her to understand the intentions and thus the intended meaning of the communicative partner (e.g., infant and caregiver are engaged in jointly attending to a toy, and the caregiver holds it out for the infants to grab, while saying *Here, the ball!*, allowing the infant to infer that *ball* must refer to the object being offered). While successful at explaining some aspects of language learning, e.g., word learning, this approach has not been able to offer an account of how (complex) syntactic structure might be learned.

More generally, while the empiricist approaches have indeed offered valuable insight into the mechanisms underlying learning tasks such as lexical acquisition or the development of semantics, they remain relatively uninformative about how syntactic structure may be learned.

4. Empirical findings inspired by generative linguistics

This mentalist and biologically based account of language acquisition proposed by generativists gave rise to a new research agenda, drawing attention to some previously unstudied phenomena and revealing links between observations hitherto believed to be unconnected. First, assuming the existence of innate mental contents pushed researchers to investigate very young infants, including newborns and even fetuses. While the late emer-

gence of an ability is not necessarily an argument against its innateness (i.e., teeth or sexual maturation), its early presence, i.e., before experience and opportunities for learning occur, is strong evidence in favor of it. Prelinguistic infants' cognitive and linguistic abilities thus became a major focus of interest. Second, and relatedly, since observable behavior was no longer the only admissible source of data, studies on perception multiplied. This was particularly relevant for studying prelinguistic infants, as their overt behavioral repertoire is strongly limited and linguistic production does not begin before the first birthday. The realization that prelinguistic infants nevertheless possess considerable knowledge about speech and language would not have been possible without investigating perception. Third, the idea that language is a biological phenomenon brought about an evolutionary perspective, with an increasing number of studies comparing the abilities of humans and animals. Several types of non-human animals provide relevant comparisons. Primates are interesting as they are our closest relatives and thus share a large part of our evolutionary history and genetic endowment. Birds are often studied because they are vocal communicators and as such have sophisticated vocalizations, perceptual and vocal learning abilities. Dogs have also been increasingly compared to humans, as they share our cultural history through domestication, and "accept" humans in their social structure. Non-human animals are often compared to human adults as well as to human infants. This latter comparison is all the more relevant, as both animals and human infants are non-linguistic, providing a particularly close match in terms of linguistic experience and knowledge, as well as in the experimental methods applicable to them. Fourth, as another consequence of this biological perspective, neurological and brain research has also been brought to bear on speech and language processing and even more abstractly on formal linguistic theories, especially with the recent advent of brain imaging methods (which, of course, in itself is not brought about my the generativist stance).

The empirical data available on language development has thus radically increased and diversified in the last decades, partly due to the cognitive and biologically-based perspective of the generative enterprise. Below, we review some of the most important empirical findings, asking whether and if yes how they confirmed or infirmed the original generativist assumptions about language.

4.1. Evidence for the biological basis of language from language acquisition

Research supporting the idea that language is part of our biology, and thus obeys principles that other biological systems also show started very early after the cognitive revolution. One of the most important findings, proposed first by Lenneberg (1967), has been that language is a critical period phenomenon: during the first years of life, language input allows a child to become a native speaker of the language(s) she is exposed to. After the closure of this window of opportunity, learners may still develop high proficiency, but will most typically not achieve full native competence.

Early evidence for this proposal came from different sources. First, several cases of linguistic isolates or feral children were found, who practically received no language input due to parental neglect. Among these children Genie and Chelsea, discovered at ages 13 and 31 years, respectively, never developed native competence after they started receiving language input, while Isabelle, who was found at age 6, rapidly caught up with her peers and learned English natively (Curtiss et al. 1974; Snow 1987). One crucial difference that may explain the different outcomes in these three cases is the age at which these individuals started receiving typical language input. Only Isabelle, the child who started learning language before the onset of puberty, could acquire language natively. Similar observations have been made about second language learners (Snow & Hoefnagel-Hohle 1978; Johnson & Newport 1989). Immigrants to the USA who arrived between ages 3–7 years developed fluency in English that was undistinguishable from that of infants born and raised in English-speaking American families, while immigrants who arrived after this age showed competence that decreased gradually with increasing age of arrival. These observations confirm that language is a critical or sensitive period phenomenon, similarly to other well-documented biological phenomena such as the development of ocular dominance in the visual cortex or imprinting in young animals.

The discovery of a critical period for language placed even more emphasis on the importance of acquisition for the understanding of human language. Studies on early development, especially on children's early production in different languages began, documenting the fact that across languages, young children go through approximately the same language developmental milestones. Babbling starts around 4-6 months, the production of the first words at 12 months, of the first two-word combinations at 2 years, and "telegraphic speech", i.e., combinations of several content words with some of the functors omitted, at 3 years. The fact that young children from different cultures follow the same developmental trajectory, despite

important differences between the grammatical and vocabulary structures of the target languages, was also taken as evidence for the biological roots of language.

An even more striking demonstration came from Nicaraguan sign language (Senghas & Coppola 2001; Kegl 2002). Deaf children born to hearing/speaking families typically do not receive language input as they do not hear the speech around them, and they develop simple signs to communicate, called home signs. In the 70s–80s in Nicaragua, such home signer children were brought together in special schools. In the span of a few decades, generations of these children developed a new, full-blown language, known as the Nicaraguan Sign Language. Remarkably, this language emerged without any input from an existing natural language, spoken or signed, as the children initially did not have full language competence, only using home sign. The fast and effortless emergence of this new sign language is considered strong evidence that the human brain is born "language-ready" (Kegl 2002), biologically prepared to process language, and is able to create it, even in the face of absent or fragmentary input.

Research on the language development of blind children led to similar conclusions. While these children are deprived of a very important source of perceptual input about the world, and about the possible meaning of linguistic communication around them, their language development is largely unperturbed (Landau et al. 1985). This suggests again that language development is to a large extent independent of perceptual experience and is driven by internal biological constraints.

4.2. Research on early speech perception and language development

The above-discussed general biological principles underlying language have highlighted the importance of early language development. Infants' and young children's speech perception and language learning abilities have started receiving much scrutiny, especially once experimental methods to test infants under laboratory circumstances were developed in the 70s–80s. We discuss some of the most interesting and relevant empirical findings below. It needs to be noted that this is far from being an exhaustive review.

4.2.1. Becoming a native listener: perceptual narrowing and phonological development

The critical period for language is best documented and closes earliest in the phonological domain (Werker & Hensch 2015). Newborns and young

infants are able to make almost all phonological discriminations that exist in the world's languages, including those that are not found in the language(s) they are learning (Eimas et al. 1971). After several months of experience with the native language, sensitivity to non-native phonological distinctions, including phoneme, tone and lexical stress contrasts, decreases, and by the end of the second half of the first year of life, infants become better at discriminating the contrasts found in their native language, while unable to distinguish most, albeit not all, non-native contrasts (Werker & Tees 1984; Kuhl 1993; Best 1994; Mattock et al. 2008; Narayan et al. in press). Importantly, during this process, the initial perceptual boundaries in phonological space are suppressed, maintained or sharpened as a function of experience, but entirely new boundaries are not created.

Perceptual reorganization is accompanied by an increased brain specialization for the native language, i.e., an increasing left-lateralization, and a gradual decrease in the neural plasticity of the brain areas involved in speech perception (Sato et al. 2010; Minagawa-Kawai et al. 2007; Friederici et al. 2007; Friederici 2011).

Similar narrowing is also observed in other perceptual domains, such as face perception, anchoring speech perception development in more general neural maturational processes (Maurer & Werker 2014; Slater et al. 2010; Watson et al. 2014).

4.2.2. Children go beyond the input in their production: evidence for abstract structural representations

The hypothesis of an innate universal grammar predicts that children should have abstract grammatical representations, which help them learn grammar. Considerable empirical evidence has accumulated suggesting that young children are indeed able to go beyond imitating the input they receive and organize their linguistic knowledge into structure-sensitive, abstract rules from their earliest productions.

One classical demonstration comes from an experimental paradigm known as the *wug*-test (Berko 1958). Preschool children were shown picture cards in which a cartoon-like animal or an action performed by a person was named using a nonsense word, obeying English phonotactic regularities. The children were then prompted to generate the plural or the past tense of the nonsense "nouns" and "verbs" (e.g., *This is a wug. Now there is another one.* [...] *There are two __.*) Children successfully created the expected forms (*wugs*) despite the fact that these were nonsense words, which they never heard before. This confirms that young children

have a representation of morphological rules that is sufficiently abstract to allow generalization to a new item.

Similar results were obtained for the existence of structure-sensitive rules in syntax (Crain & Nakayama 1987). Preschoolers, who were familiar with the formation of yes/no questions by subject/auxiliary inversion in English (*The man is tall.* → *Is the man ___ tall?*), were given a sentence elicitation task in which they had to form yes/no question from complex sentences with a subject relative (*The man who is tall is in the other room.*), a construction they do not yet produce and which occurs infrequently in the input. There are infinitely many possible generalizations that are compatible with the simple yes/no question (e.g., front the leftmost auxiliary, front the main clause auxiliary etc.). However, young children seem to only entertain the structure-sensitive (and correct) one (front the main clause auxiliary: *Is the man who is in the other room ___ tall?*), not the others, as they do not make mistakes of the type *Is the man who ___ in the other room is tall.*

While even very young infants readily generalize beyond the input, their early production is not error-free. Interestingly, many of the errors they make also provide evidence for the application of abstract rules: children often overgeneralize. Forms such as **goed* (instead of *went*) are not uncommon. Since these forms are not produced by adults, and are thus absent form the input, they provide clear evidence for the existence of abstract rules in children's early grammars. Young children's multiword speech thus provides evidence for abstract representations in production from the toddler years onwards. Can we find even earlier evidence? Since there is no considerable production before this age, we need to turn to perception and consider whether infants' speech perception and language learning abilities provide any cues. The artificial grammar learning paradigm was adapted to infants with this purpose.

4.2.3. Artificial grammars: a central paradigm

Originally created to study implicit sequence learning (Reber 1967; 1969), the artificial grammar learning paradigm was adapted to investigate language learning in adults and infants. In this paradigm, a set of rules are defined over a vocabulary of typically nonsense (or a mix of nonsense and existent) words. Learners are exposed to grammatical strings generated by (some of) the rules of the artificial grammar, and are tested on the learned strings, new grammatical strings (generalization) and/or new ungrammatical strings. This paradigm allows researchers to have more control over the stimulus material than if they used stimuli from a natural language, and

they can, to a large extent, reduce the effects of experience. This paradigm
is thus suitable to emulate language acquisition in adult learners and it is
also attractive to investigate early language acquisition, because the stim-
ulus material can be simplified to make up for infants' attentional and
memory limitations.

This paradigm has been widely used (for reviews, see Gomez & Gerken
2000; Gervain et al. 2018) to study statistical learning, lexical acquisition,
as well as the acquisition of morphological, syntactic and semantic reg-
ularities in infants and young children (e.g., Saffran et al. 1996; Marcus
et al. 1999; Dawson & Gerken 2009; Gomez & Gerken 1999; Marchetto &
Bonatti 2015).

A seminal artificial grammar learning study (Marcus et al. 1999)
aimed to show that very young infants are able to learn abstract, "al-
gebraic" rules, i.e., rules containing variables. The identity relation was
chosen as the target rule as it is the simplest abstract relation possible.
Infants were trained on a grammar generating three-word strings with two
identical and one different item (ABB, AAB or ABA), and they were
tested on their ability to discriminate the trained grammar from the other
grammars. For instance, infants familiarized with AAB sequences were
tested with new AAB sequences, i.e., sequences consistent **with** the gram-
mar of familiarization, as well as on ABA sequences, inconsistent with the
familiarization grammar. Infants discriminated the inconsistent sequences
from the consistent ones. Importantly, infants showed discrimination for
sequences made up of words that were not presented during familiariza-
tion (e.g., if familiarization items were "ba po ba", "ko ga ko", "ba ba po",
"ko ko ga", test items were "wo fe fe", etc.). This way, it could be shown
that infants generalize the underlying abstract rule, rather than relying
on item-based information, e.g., statistics (frequency of occurrence or co-
occurrence, etc.) or the specific position of a given string element.

These initial results gave rise to a large body of literature further inves-
tigating how infants process, learn and represent identity-based structures,
whether this processing is specific to the language domain and to what ex-
tent it can be considered abstract (Marcus et al. 2007; Dawson & Gerken
2009; Saffran et al. 2007; Frank et al. 2009; Johnson et al. 2009). Most im-
portantly for our purposes, the interpretation that these simple repetition-
based structures are represented as algebraic rules was questioned. It has
been argued that adjacent repetitions are salient, Gestalt-like primitives,
which may be detected automatically by the perceptual system without
recourse to symbols or abstract representations (Endress et al. 2009). How-
ever, even if the detection of identity/repetition is a primitive, explaining

the ability to discriminate adjacent repetitions (ABB) from non-adjacent ones (ABA, as in Marcus et al. 1999) or random sequences (ABC, as in Gervain et al. 2008; Gervain & Werker 2012), this account cannot explain how infants discriminate between two adjacent repetition based structures, such as AAB vs. ABB, an ability observed in 7-month-olds (Marcus et al. 1999) and newborns (Gervain et al. 2012). This is because discriminating AAB and ABB structures requires the combination of repetition detection with sequential ordering (initial vs. final position of the repetition).

Artificial grammar learning is thus a paradigm that has proven particularly useful in exploring the nature of infants' knowledge and representations of language structure.

4.2.4. Bootstrapping approaches to the acquisition of grammar

The above-presented empirical results probe the acquisition of specific aspects of grammar and vocabulary. But how do they scale up to explain the complex process of language acquisition? As discussed above, the P&P paradigm provides a general approach to explain language acquisition universally across all languages. However, the linking problem arises (Pinker 1984): how do infants know what cues in the input to pay attention to in order to set a given parameter? Bootstrapping mechanisms rely on the assumption that infants can organize their linguistic knowledge by analysing the perceptually available cues present in the input that, in turn, correlate with the underlying linguistic structure.

Pinker (1984) was one of the first to introduce the notion of "bootstrapping". He argued for an innate correspondence between semantics and syntax. According to his **semantic bootstrapping hypothesis**, the basic semantic notions present in our every-day life (e.g., "actions" or "concrete objects") can be linked to syntactic or lexical categories (e.g., "verbs" and "nouns"). Infants may rely on this innately specified knowledge to expect nouns to refer to objects, and verbs to refer to actions. By observing the contingency between specific words and their meanings, infants might be able to link semantic information to the syntactic structure of a sentence (**linking rules**).

Others have argued that this semantic bootstrapping is not always possible, as infants would need to be able to process sentence structure to begin with in order to be able to establish the correspondence with semantics. The **syntactic bootstrapping hypothesis** (Gleitman 1990) proposes that learning goes the other way: it is by observing the syntactic structure that infants may deduce knowledge of the meaning. Specifically, this procedure "deduces the word meanings from the semantically relevant

syntactic structures associated with a verb in input utterances" (*ibid.*, 30).
For instance, a transitive verb is likely interpreted as an action performed
by an agent on an object (e.g., *he sent a letter*), whereas an intransitive
verb is more typically an action performed by an agent with no object
involved (e.g., *she was laughing*). Using this regularity, infants may cate-
gorize a verb as transitive if it appears in the sentence with two nominal
arguments, but as intransitive if only one noun phrase is present. This
is a plausible learning mechanism to acquire verbs, because infants learn
nouns earlier than verbs. Indeed, infants' first 50–100 words mainly consist
of nouns.

Importantly, infants are also sensitive to the acoustic information car-
ried by the speech stream from very early on. It has therefore been sug-
gested that prosodic and phonological cues contained in speech help infants
extract information about grammatical structure. The central argument of
the **prosodic/phonological bootstrapping hypothesis** (Gleitman & Wanner
1982; Nespor et al. 2008; Morgan & Demuth 1996) suggests that certain
acoustic/phonological properties of speech cue structural properties of syn-
tax. Variations in duration, intensity and pitch in the speech signal are
systematically related to the prosodic hierarchy, which in turn correlates
with the syntactic hierarchy of the sentence (Nespor & Vogel 1986; Selkirk
1984; Nespor et al. 2008). Therefore, these acoustic cues might help infants
parse speech into smaller, syntactically relevant units (Morgan & Demuth
1996; Christophe et al. 1997; 2003).

One domain in which prosodic bootstrapping has been well estab-
lished is the acquisition of word order. Languages of the world vary sys-
tematically in the relative order of their principal syntactic components
(Greenberg 1978; Dryer 1992), such as the Verb (V) and its Object (O)
and more generally functors and their corresponding content words. Im-
portantly from the point of view of language acquisition, the two possi-
ble word orders, VO/functor-initial and OV/functor-final, have different
prosodic and acoustic correlates (Nespor & Vogel 1986; Nespor et al. 2008;
Christophe et al. 2003; Gervain & Werker 2013). The phonological phrase
in VO languages (e.g., French, English and Italian) is characterized by
final prominence, marked by increased duration on the prominent item
as compared to the non-prominent item (e.g., *to **Ro:me***). OV languages
(e.g., Turkish, Japanese and Basque), by contrast, are characterized by
initial prominence, marked by increased pitch or intensity on the promi-
nent as compared to the non-prominent element (e.g., Japanese: *^**Tokyo**
kara* Tokyo from 'from Tokyo'). Languages with different word orders thus
use different acoustic cues to mark prominence. Acoustic cues are readily

available in the speech signal, and a well established auditory bias, known as the Iambic–Trochaic Law (Hayes 1995), suggests that sound sequences in which elements contrast in duration are naturally perceived as forming units with an iambic, prominence-final pattern, while sequences in which elements contrast in pitch or intensity are perceived trochaically, i.e., with initial prominence. This correlates perfectly with the underlying prosodic and syntactic structures, and thus offers infants a readily available, perceptual, language-independent cue to word order, one of the most basic properties of grammar. Experimental evidence has indeed demonstrated that infants readily rely on this acoustic/prosodic cue. Newborn infants show a preference for the acoustic pattern (iambic for durational contrasts and trochaic for pitch/intensity contrasts) found in the language(s) heard prenatally (Abboub et al. 2016) and they are also sensitive to the phonological differences between functors and content words (Shi et al. 1999). Infants between 6 and 12 weeks of age are able to distinguish between the prosodic patterns of sentences in French, a VO language, and Turkish, an OV language, on the basis of these acoustic differences alone, i.e., even when sentences are delexicalized by replacing all vowels with a schwa and all consonants with a pre-defined member of their respective manner of articulation category. Furthermore, 8-month-old OV-VO bilinguals can use these acoustic cues to guide their choice of word order in an artificial grammar task (Gervain & Werker 2013).

Another well-established cue that can be used in combination with phrasal prosody to learn word order is word frequency, because prosodic and word frequency information are aligned at the phrasal level. In VO languages, prosodic prominence falls on the final constituent of the phrase, which is typically a content word, whereas its functors, which precede it, are non-prominent. In OV languages, prominence also falls on the content word, but in these languages content words tend to be phrase-initial. The **frequency-based bootstrapping hypothesis** (Gervain et al. 2008; Bernard & Gervain 2012) relies on the language universal division between functors and content words (Chomsky 1995; Fukui 1986; Abney 1987). The two classes differ in their frequency of occurrence. Individual functors have a much higher frequency of occurrence than individual content words (Cutler & Carter 1987; Kucera & Francis 1967; Gervain et al. 2008). Sensitivity to this difference in word frequency distribution is detectable pre-lexically. Eight-month-olds exposed to languages with opposite word orders, e.g., functor-initial Italian and functor-final Japanese, showed opposite preferences for word order in an artificial grammar task. Italian infants preferred sequences starting with a frequent word, while Japanese infants preferred

sequences starting with an infrequent word, mirroring the word orders of these two languages (Gervain et al. 2008).

Word frequency is also used as a cue to establish the categories of function and content words. Content words come in open classes (e.g., *iPad, Brexit* etc.), whereas functors constitute closed classes, into which no new items can be added without a major language change. At 8 months, infants treat frequent words as belonging to closed classes, accepting no replacement for the frequent words in an artificial grammar, whereas they process infrequent words as belonging to open classes readily accepting substitutions with new items (Marino, Bernard, and Gervain under review). As a confirmation, 17-month-old infants expect infrequent, but not frequent words to have semantic content, serving as possible labels for objects (Hochmann et al. 2010). This further strengthens the claim that frequency acts as a cue to true lexical categories. Interestingly, non-linguistic animals (rats) are also sensitive to frequency information in a similar artificial grammar paradigm, but do not encode the relative position of frequent and infrequent words (Toro et al. 2016).

The prosodic and frequency-based bootstrapping hypotheses are particularly attractive, because they provide perceptually based mechanisms to explain how infants might start breaking into their native grammar(s) very early, before and thus independently of the lexicon and implies the existence of a rudimentary representation of word order as early as 8 months of age.

In sum, a substantial amount of empirical evidence has gathered suggesting that young infants, even newborns, exhibit more sophisticated speech perception and language learning abilities, as well as more knowledge about the specific properties of their native language than behaviorist approaches prior to the cognitive turn would have predicted. Specific mechanisms have been proposed to account for language acquisition cross-linguistically in line with formal theories of language.

4.3. Comparative research

The evolution of language has received considerable attention (e.g., Bickerton 2016; Pinker & Bloom 1990; Hauser et al. 2002; Christiansen & Kirby 2003), especially since the biological nature of language came to the forefront. Historical evidence in this domain is scare. In the last 15–20 years, however, cross-species research directly comparing humans' and animals' performance in language learning tasks has provided invaluable insight into the abilities we share with non-human animals, and those that are unique

to humans, possibly to human language. This has been particularly productive in the artificial grammar learning and neuroimaging domains, in which adults, infants and animals can easily be compared, as no prior linguistic knowledge is necessary to complete the experimental tasks. In a seminal, but much debated paper, Hauser et al. (2002) suggested that the only unique human evolutionary step towards language is the emergence of the combinatorial operation ("merge" in the sense of Chomsky 1995), underlying humans' ability to entertain hierarchically embedded, recursive representations, unavailable to non-human animals. They backed this claim up with comparative artificial grammar learning data from monkeys and humans (Fitch & Hauser 2004). Several subsequent publications suggested that some animal species, e.g., songbirds, are able to learn centrally embedded, recursive structures, similarly to humans (Gentner et al. 2006), while others argued that in the original tasks, even humans did not necessarily learn such complex structures (Hochmann et al. 2008). A large amount of recent work (for a recent summary, see the special issue, Petkov & Marslen-Wilson 2018) has used the same artificial grammar learning paradigms as those used with infants to test whether animals are able to extract identity-/repetition-based regularities (in primates: Rey et al. 2019; in song birds: Chen et al. 2015), non-adjacent dependencies (in primates: Malassis et al. 2018; Ravignani et al. 2013; Wilson et al. 2013) or prosodic patterns (in rats: de la Mora et al. 2012; in song birds: Spierings et al. 2017). The debate is still open as to whether animals have recursive linguistic representations and about the level of the Chomsky hierarchy that best characterizes their competence.

4.4. Neuroimaging studies

In the past decades, brain imaging techniques have revolutionized our ability to explore the human brain. How the brain processes speech and language has been investigated in thousands of electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS) and positron emission tomography (PET) studies from birth to adulthood. The brain network responsible for language processing is now quite well described in adults (e.g., Friederici 2011; Poeppel 2014; Kutas and van Petten 1988; Bastiaansen & Hagoort 2006) and its origins are starting to be understood in infants (Dehaene-Lambertz et al. 2002; 2008; Gervain et al. 2011; Friederici et al. 2007; Pena et al. 2003).

While this technological development and knowledge of brain function and anatomy it brought about are independent of the generative tradition, the relevance and interpretation of brain imaging data for language sciences have definitively been influenced by the biological stance of the generative enterprise. If language is an "organ", brain imaging should allow us to investigate it. Indeed, certain imaging studies were directly inspired by linguistic theory and explicitly set out to find evidence for the neural mechanisms encoding formal linguistic operations, such as merge or move, and possible vs. impossible linguistic rules (Friederici et al. 2003; Friederici & Frisch 2000; Musso et al. 2003). Thus, Musso and colleagues have shown increased activation in Broca's area when German-speakers syntactic rules of Japanese or Italian, but no activation when the same Japanese or Italian words were used in rules that are not allowed by UG, such as counting rules. Importantly for the issue of language development, evidence has been found that the infant language network is already similar to the mature, adult network. In particular, both the temporal, auditory areas and the frontal regions, responsible for higher order structural or sequential processing, are operational from birth (Dehaene-Lambertz et al. 2002; Pena et al. 2003; Gervain et al. 2008). While many brain imaging studies were unable to identify the neural correlates of different syntactic constructions as posited by generative grammar, and even provided evidence against generativist assumptions (Pulvermüller 1999; van Turennout et al. 1998; Bastiaansen et al. 2010), the generativist quest for evidence in favor of the mental and biological reality of these theoretical constructs bolstered the idea that neuroimaging can directly inform language theory, and not just investigations of brain function, anatomy or pathology.

5. The most recent perspective: integrative models

As summarized above, the nativist-empiricist (or "nature-nurture") debate dominated much of the second half on the 20th century as competing paradigms in explaining the language faculty. As a result, we now have a much better empirical understanding about how the human mind and brain process and learn language. Very early language acquisition has played a pivotal role in these debates, and much has been learned about how babies break into language. Interestingly, the influence was reciprocal. Generative theory also changed as a result of empirical evidence, as reviewed in Chomsky (2007). Specifically, research on language acquisition, and in particular, the problem of learnability has given a strong push to-

wards simpler syntactic formalism, which Chomsky's *Minimalist Program* (Chomsky 1995) made explicit.

These empirical observations, as well as recent advances in brain imaging, genetics and epigenetics (Werker & Hensch 2015; Lewkowicz 2000), have led to new theoretical questions and a more integrated view of innate and learned abilities, rendering the strict nature-nurture dichotomy obsolete.

The last decades of biological research have in fact proposed a new model of how gene expression is regulated by inherent biological factors, but also by experience. It has been proposed that DNA expression can be modified during the entire life span of an individual through experience and environmental factors. This up- or downregulation of DNA expression is transmissible (e.g., during genome imprinting or embryonic stage), thus these epigenetic re-arrangements change gene expression without changing the DNA sequences (Roth & Sweatt 2009). Epigenetics studies these flexible re-arrangements of our genome. The human genome is no longer considered exclusively as a "committed" blueprint, but rather as a "potentiality".

Epigenetic modulations of gene activity may be induced by environmental factors such as diet, stress, exposure to certain chemical substances and specific behavioral stimulations, e.g., extensive practice, sensory deprivation etc. (e.g., Franklin & Mansuy 2010; Gräff & Mansuy 2008; Jirtle & Skinner 2007; Liu et al. 2009; Roth & Sweatt 2009; Zhang & Meaney 2010). These modulations have most often been studied in animals. One study has shown, for instance, that better quality infant-caregiver interaction (e.g., more time spent nursing) produces epigenetic modification that impact the stability of gene expression in the central nervous system, thereby modifying behavior (Roth & Sweatt 2009).

In this perspective, dynamic and complex interactions take place between the environment and the genome (Waddington 1942). These interactions trigger epigenetic modulations as a response to the experience, forming epigenetic "memories" of the external world. Since the earliest stages of the development, environmental feedback is necessary to develop optimal functioning. Behavioral outcomes are thus often mediated by the long-term influence of these experiences.

This framework puts the notions of **critical** or **sensitive periods** and brain plasticity into a new perspective (Werker & Hensch 2015). Different aspects of language acquisition have different temporal windows of opportunity. The critical periods of phonetic, lexical and syntactic learning are different (Kuhl 2010). This series of sensitive periods for language

creates a cascading dynamics for language development, where each onto-
genetic accomplishment supports and interacts with the subsequent ones.
This epigenetic cascade "can be altered by sensory deprivation, pharmaco-
logical exposure and linguistic experience" (Werker & Hensch 2015, 187).
For instance, under specific circumstances such as deafness or exposure
to drug like serotonin reuptake inhibitors (SRI), the opening and closing
of the critical period is shifted (Weikum et al. 2012). In congenitally deaf
people, measures of acoustic discrimination indicate that plasticity stays
in place for a longer period than in normally hearing people (Faulkner &
Pisoni 2013; Kral & Sharma 2012). Also, untreated maternal depression
delays the closure of the sensitive period, while exposure to SRI (in infants
whose mothers are depressed, but medicated), accelerates it (Weikum et
al. 2012).

These advances show that both genetically endowed and environ-
mental factors play an important role, and the most important ques-
tion is to understand how they interact synergistically to enable human
development.

Acknowledgments

References

Abboub, Nawal, Thierry Nazzi and Judit Gervain. 2016. Prosodic grouping at birth. Brain
    and Language 162. 46–59.

Abney, Steven Paul. 1987. English NP in its sentential aspect. Doctoral dissertation. MIT.

Bastiaansen, Marcel and Peter Hagoort. 2006. Oscillatory neuronal dynamics during lan-
    guage comprehension. Progress in Brain Research 159. 179–196.

Bastiaansen, Marcel, Lilla Magyari and Peter Hagoort. 2010. Syntactic unification oper-
    ations are reflected in oscillatory dynamics during on-line sentence comprehension.
    Journal of Cognitive Neuroscience 22. 1333–1347.

Berko, Jean. 1958. The child's learning of English morphology. Word 14. 150–177.

Bernard, Carline and Judit Gervain. 2012. Prosodic cues to word order: What level of
    representation? Frontiers in Language Sciences 3. 451.

Best, Catherine T. 1994. The emergence of native-language phonological influences in
    infants: A perceptual assimilation model. In J. C. Goodman and H. C. Nusbaum

(eds.) The development of speech perception: The transition from speech sounds to spoken words. Cambridge, MA: MIT Press. 167–224.

Bickerton, Derek. 2016. Roots of language. Berlin: Language Science Press.

Boas, Franz. 1940. Race, language and culture. University of Chicago Press: Chicago.

Braine, Martin D. S. 1971. On two types of models on the internalization of grammars. In D. I. Slobin (ed.) The ontogenesis of grammar. New York: Academic Press. 153–18.

Bybee, Joan L. and Paul Hopper (eds.). 2001. Frequency and the emergence of linguistic structure. Amsterdam & Philadelphia: John Benjamins.

Chen, Jiani, Danielle van Rossum and Carel Ten Cate. 2015. Artificial grammar learning in zebra finches and human adults: XYX versus XXY. Animal Cognition 18. 151.

Chomsky, Noam. 1957. Syntactic structures. The Hague: Mouton.

Chomsky, Noam. 1959. A review of B. F. Skinner's *Verbal behavior*. Language 35. 26–58.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Chomsky, Noam. 1995. The minimalist program. Cambridge, MA: MIT Press.

Chomsky, Noam. 2007. Approaching UG from below. In U. Sauerland and H.-M. Gärtner (eds.) Interfaces + recursion = language? Chomsky's minimalism and the view from syntax–semantics. Berlin & New York: Mouton de Gruyter. 1–29.

Christiansen, Morten H. and Simon Kirby. 2003. Language evolution. Oxford: Oxford University Press.

Christophe, Anne, Teresa Guasti and Marina Nespor. 1997. Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. Language and Cognitive Processes 12. 585–612.

Christophe, Anne, Marina Nespor, Maria Teresa Guasti and Brit Van Ooyen. 2003. Prosodic structure and syntactic acquisition: The case of the head-direction parameter. Developmental Science 6. 211–220.

Crain, Stephen. 1991. Language acquisition in the absence of experience. Behavioral and Brain Sciences 14. 597–650.

Crain, Stephen and Mineharu Nakayama. 1987. Structure dependence in grammar formation. Language 63. 522–543.

Curtiss, Susan, Victoria A. Fromkin, Stephen Krashen, David Rigler and Marilyn Rigler. 1974. The linguistic development of genie. Language 50. 528–554.

Cutler, Anne. and David M. Carter. 1987. The predominance of strong initial syllables in the English vocabulary. Computer Speech and Language 2. 133–42.

Dawson, Colin and LouAnn Gerken. 2009. From domain-generality to domain-sensitivity: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. Cognition 111. 378–382.

Dehaene-Lambertz, Ghislaine, Stanislas Dehaene and Lucie Hertz-Pannier. 2002. Functional neuroimaging of speech perception in infants. Science 298. 2013–2015.

Dehaene-Lambertz, Ghislaine, Lucie Hertz-Pannier, Jessica Dubois and Stanislas Dehaene. 2008. How does early brain organization promote language acquisition in humans? European Review 399. 399–411.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. Language 68. 81–138.

Eimas, Peter D., Einar R. Siqueland, Peter W. Jusczyk and James Vigorito. 1971. Speech perception in infants. Science 171. 303–306.

Elman, Jeff, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett. 1996. Rethinking innateness: A connectionist perspective on development. Cambridge, MA: MIT Press.

Endress, Ansgar D., Marina Nespor and Jacques Mehler. 2009. Perceptual and memory constraints on language acquisition. Trends in Cognitive Sciences 13. 348–353.

Faulkner, Kathlene F. and David B. Pisoni. 2013. Some observations about cochlear implants: Challenges and future directions. Neuroscience Discovery 1. 9.

Fitch, W. Tecumseh and M. D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. Science 303. 377–380.

Frank, Michael C., Jonathan A. Slemmer, Gary F. Marcus and Scott P. Johnson. 2009. Information from multiple modalities helps five-month-olds learn abstract rules. Developmental Science 12. 504–509.

Franklin, Tamara B. and Isabelle M. Mansuy. 2010. Epigenetic inheritance in mammals: Evidence for the impact of adverse environmental effects. Neurobiology of disease 39. 61–65.

Friederici, Angela D. 2011. The brain basis of language processing: From structure to function. Physiological Reviews 91. 1357–1392.

Friederici, Angela D., Manuela Friedrich and Anne Christophe. 2007. Brain responses in 4-month-old infants are already language specific. Current Biology 17. 1208–1211.

Friederici, Angela D. and Stefan A. Frisch. 2000. Verb argument structure processing: The role of verb-specific and argument-specific information. Journal of Memory and Language 43. 476–507.

Friederici, Angela D., Shirley-Anne Ruschemeyer, Anja Hahne and Christian J. Fiebach. 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. Cerebral Cortex 13. 170–177.

Fukui, Naoki. 1986. A theory of category projection and its applications. Doctoral dissertation. MIT.

Gentner, Timothy Q., Kimberly Fenn, Daniel Margoliash and Howard Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. Nature 44. 1204–1207.

Gervain, Judit, Iris Berent and Janet F. Werker. 2012. Binding at birth: The newborn brain detects identity relations and sequential position in speech. Journal of Cognitive Neuroscience 24. 564–574.

Gervain, Judit, Irene de la Cruz Pavia and LouAnn Gerken. 2018. Behavioural and imaging studies of infant artificial grammar learning. Topics in Cognitive Science. https://doi.org/10.1111/tops.12400.

Gervain, Judit, Jacques Mehler, Janet F. Werker, Charles A. Nelson, Gergely Csibra, Sarah Lloyd-Fox, Mohinish Shukla and Richard N. Aslin. 2011. Near-infrared spectroscopy: A report from the McDonnell Infant Methodology Consortium. Developmental Cognitive Neuroscience 1. 22–46.

Gervain, Judit, Marina Nespor, Reiko Mazuka, Ryota Horie and Jacques Mehler. 2008. Bootstrapping word order in prelexical infants: A Japanese–Italian cross-linguistic study. Cognitive psychology 57. 56–74.

Gervain, Judit and Janet F. Werker. 2012. Learning non-adjacent regularities at age 0;7. Journal of Child Language FirstView. 1–13.

Gervain, Judit and Janet F. Werker. 2013. Prosody cues word order in 7-month-old bilingual infants. Nature Communications 4. 1490.

Gleitman, Lila. 1990. The structural sources of verb meanings. Language Acquisition 1. 3–55.

Gleitman, Lila R. and Barbara Landau. 1994. The acquisition of the lexicon. Cambridge, MA: MIT Press.

Gleitman, Lila R. and Eeric Wanner. 1982. Language acquisition: The state of the state of the art. In E. Wanner and L. R. Gleitman (eds.) Language acquisition: The state of the art. Cambridge, MA: MIT Press. 3–48.

Gomez, R. L. and L. Gerken. 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70. 109–135.

Gomez, R. L. and L. Gerken. 2000. Infant artificial language learning and language acquisition. Trends in Cognitive Sciences 4. 178–186.

Gräff, Johannes and Isabelle M. Mansuy. 2008. Epigenetic codes in cognition and behaviour. Behavioural Brain Research 192. 70–87.

Greenberg, Joseph H. (ed.). 1978. Universals of human language. Stanford, CA: Stanford University Press.

Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? Science 298. 1569–1579.

Hochmann, Jean-Remy, Ansgar D. Endress and Jacques Mehler. 2010. Word frequency as a cue for identifying function words in infancy. Cognition 115. 444–457.

Hochmann, Jean-Remy, Azadpour Mahan and Mehler Jacques. 2008. Do humans really learn AnBn artificial grammars from exemplars? Cognitive Science 32. 1021–1036.

Jirtle, Randy L. and Michael K. Skinner. 2007. Environmental epigenomics and disease susceptibility. Nature reviews genetics 8. 253–262.

Johnson, Jacqueline S. and Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. Cognitive Psychology 21. 60–100.

Johnson, Scott P., Keith J. Fernandes, Michael C. Frank, Natasha Kirkham, Gary Marcus, Hugh Rabagliati and Jonathan A. Slemmer. 2009. Abstract rule learning for visual sequences in 8-and 11-month-olds. Infancy 14. 2–18.

Kegl, Judy. 2002. Language emergence in a language-ready brain. Directions in Sign Language Acquisition 2. 207.

Kral, Andrej and Anu Sharma. 2012. Developmental neuroplasticity after cochlear implantation. Trends in Neurosciences 35. 111–122.

Kucera, Henry and W. Nelson Francis. 1967. Computational analysis of present-day American English. Providence: Brown University Press.

Kuhl, Patricia K. 1993. Innate predispositions and the effects of experience in speech perception: The Native Language Magnet Theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage and J. Morton (eds.) Developmental neurocognition: Speech and face processing in the first year of life. Dordrecht: Springer. 259–274.

Kuhl, Patricia K. 2010. Brain mechanisms in early language acquisition. Neuron 67. 713–727.

Kutas, Marta and Cyma K. van Petten. 1988. Event-related brain potential studies of language. Advances in Psychophysiology 3. 139–187.

Landau, Barbara, Lila R. Gleitman and Barbara Landau. 1985. Language and experience: Evidence from the blind child. Cambridge, MA: Harvard University Press.

Lany, Jill and Jenny R. Saffran. 2010. From statistics to meaning: Infants' acquisition of lexical categories. Psychological Science 21. 284–291.

LeCun, Yann, Joshua Bengio and Geoffrey Hinton. 2015. Deep learning. Nature 521. 436–444.

Lenneberg, Eric H. (ed.). 1967. Biological foundations of language. New York: Wiley.

Lewkowicz, David J. 2000. The development of intersensory temporal perception: An epigenetic systems/limitations view. Psychological Bulletin 126. 281–308.

Lidz, Jeffrey, Henry Gleitman and Lila R. Gleitman. 2003. Understanding how input matters: Verb learning and the footprint of Universal Grammar. Cognition 87. 151–178.

Liu, Liang, Thomas van Groen, Inga Kadish and Trygve O. Tollefsbol. 2009. DNA methylation impacts on learning and memory in aging. Neurobiology of Aging 30. 549–556.

MacWhinney, Brian. 2000. The CHILDES Project: Tools for analyzing talk (Third edition). Mahwah, NJ: Lawrence Erlbaum.

Malassis, Raphaëlle, Arnaud Rey and Joël Fagot. 2018. Non-adjacent dependencies processing in human and non-human primates. Cognitive Science 42. 1677–1699.

Marchetto, Erika and Luca L. Bonatti. 2015. Finding words and word structure in artificial speech: The development of infants' sensitivity to morphosyntactic regularities. Journal of Child Language 42. 873–902.

Marcus, Gary F. 1993. Negative evidence in language acquisition. Cognition 46. 53–85.

Marcus, Gary F. 2018. Deep learning: A critical appraisal. arXiv 1801.0063.

Marcus, Gary F., Keith J. Fernandes and Scott P. Johnson. 2007. Infant rule learning facilitated by speech. Psychological Science 18. 387.

Marcus, Gary F., Sugumaran Vijayan, Shoba Bandi Rao and Peter M. Vishton. 1999. Rule learning by seven-month-old infants. Science 283. 77–80.

Mattock, Karen, Monika Molnar, Linda Polka and Denis Burnham. 2008. The developmental course of lexical tone perception in the first year of life. Cognition 106. 1367–1381.

Maurer, Daphne and Janet F. Werker. 2014. Perceptual narrowing during infancy: A comparison of language and faces. Developmental Psychobiology 56. 154–178.

Minagawa-Kawai, Yasuyo, Koichi Mori, Nozomi Naoi and Shozo Kojima. 2007. Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. Journal of Neuroscience 27. 315–321.

Mora, Daniela M. de la, Marina Nespor and Juan M. Toro. 2012. Do humans and non-human animals share the grouping principles of the iambic–trochaic law? Attention, Perception, & Psychophysics 75. 1–9.

Morgan, James L. and Katherine Demuth. 1996. Signal to syntax: Bootstrapping from speech to grammar in early acquisition. Hillsdale, NJ: Lawrence Erlbaum.

Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel and Cornelius Weiller. 2003. Broca's Area and the language instinct. Nature Neuroscience 6. 774–781.

Narayan, Chandan R., Janet F. Werker and Patrice Speeter Beddor. in press. Acoustic salience affects speech perception in infancy: Evidence from nasal place discrimination. Developmental Science.

Nespor, Marina. 1990. On the rhythm parameter in phonology. In I. Roca (ed.) Logical issues in language acqisition. Dordrecht: Foris. 157–175.

Nespor, Marina, Mohinish Shukla, Ruben van de Vijver, Cinzia Avesani, Hanna Schraudolf and Caterina Donati. 2008. Different phrasal prominence realization in VO and OV languages. Lingue e Linguaggio 7. 1–28.

Nespor, Marina and Irene Vogel. 1986. Prosodic phonology. Dordrecht: Foris.

Pelucchi, Bruna, Jessica F. Hay and Jenny R. Saffran. 2009. Learning in reverse: Eight-month-old infants track backward transitional probabilities. Cognition 113. 244–47.

Pena, Marcela, Atsushi Maki, Damir Kovacic, Ghislaine Dehaene-Lambertz, Hideaki Koizumi, Furio Bouquet and Jacques Mehler. 2003. Sounds and silence: An optical topography study of language recognition at birth. PNAS 100. 11702–11705.

Petkov, Christopher I. and William D. Marslen-Wilson. 2018. Editorial overview: The evolution of language as a neurobiological system. Current Opinion in Behavioral Sciences, The Evolution of Language 21. v–xii.

Pinker, Steven. 1984. Language learnability and language development. Cambridge, MA: Harvard University Press.

Pinker, Steven and P. Bloom. 1990. Natural language and natural selection. Behavioral and Brain Sciences 13. 707–784.

Poeppel, David. 2014. The neuroanatomic and neurophysiological infrastructure for speech and language. Current Opinion in Neurobiology 28. 142–149.

Pullum, Geoffrey K. and Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. The Linguistic Review 19. 9–50.

Pulvermüller, Friedemann. 1999. Words in the brain's language. Behavioral and Brain Sciences 22. 253–336.

Ravignani, Andrea, Ruth-Sophie Sonnweber, Nina Stobbe and W. Tecumseh Fitch. 2013. Action at a distance: Dependency sensitivity in a new world primate. Biology Letters 9. 20130852.

Reber, Arthur S. 1967. Implicit learning of synthetic languages: The role of instructional set. Journal of Experimental Psychology: Human Learning and Memory 2. 88–94.

Reber, Arthur S. 1969. Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology 81. 115–119.

Rey, Arnaud, Laure Minier, Raphaëlle Malassis, Louisa Bogaerts and Joël Fagot. 2019. Regularity extraction across species: Associative learning mechanisms shared by human and non-human primates. Topics in Cognitive Science 11. 573–586.

Roth, Tania L. and J. David Sweatt. 2009. Regulation of chromatin structure in memory formation. Current Opinion in Neurobiology 19. 336–342.

Saffran, Jenny R., R. N. Aslin and E. L. Newport. 1996. Statistical learning by 8-month-old infants. Science 274. 1926–1928.

Saffran, Jenny R., Rebecca L. Seibel, Seth D. Pollak and Anna Shkolnik. 2007. Dog is a dog is a dog. Infant rule learning is not specific to language. Cognition 105. 669–680.

Sato, Yutaka, Yuko Sogabe and Reiko Mazuka. 2010. Development of hemispheric specialization for lexical pitch-accent in Japanese infants. Journal of Cognitive Neuroscience 22. 2503–2513.

Selkirk, Elisabeth O. 1984. Phonology and syntax. The relation between sound and structure. Cambridge, MA: MIT Press.

Senghas, Ann and Marie Coppola. 2001. Children creating language: How Nicaraguan sign language acquired a spatial grammar. Psychological Science 12. 323–328.

Shi, Rushen, Janet F. Werker and James L. Morgan. 1999. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. Cognition 72. B11–B21.

Skinner, Burhus F. 1957. Verbal behavior. New York: Appleton.

Slater, Alan, Paul C. Quinn, David J. Kelly, Kang Lee, Christopher A. Longmore, Paula R. McDonald and Olivier Pascalis. 2010. The shaping of the face space in early infancy: Becoming a native face processor. Child Development Perspectives 4. 205–211.

Snow, Catherine E. 1987. Relevance of the notion of a Critical Period to Language Acquisition. In M. H. Bornstein (ed.) Sensitive periods in development: Interdisciplinary perspectives. Hillsdale, NJ: Lawrence Erlbaum. 183–209.

Snow, Catherine E. and Marian Hoefnagel-Hohle. 1978. The critical period for language acquisition: Evidence for second language learning. Child Development 49. 1114–1128.

Spierings, Michelle, Jeroen Hubert and Carel ten Cate. 2017. Selective auditory grouping by zebra finches: Testing the Iambic–Trochaic Law. Animal Cognition 20. 665–675.

Teinonen, Tuomas, Vineta Fellman, Risto Näätänen, Paavo Alku and Minna Huotilainen. 2009. Statistical language learning in neonates revealed by event-related brain potentials. BMC Neuroscience 10. 21.

Tomasello, Michael. 2000. Do young children have adult syntactic competence? Cognition 74. 209–253.

Toro, Juan M., Marina Nespor and Judit Gervain. 2016. Frequency-based organization of speech sequences in a nonhuman animal. Cognition 146. 1–7.

Turennout, Miranda van, Peter Hagoort and Colin M. Brown. 1998. Brain activity during speaking: From syntax to phonology in 40 milliseconds. Science 280. 572–574.

Waddington, Conrad H. 1942. Canalization of development and the inheritance of acquired characters. Nature 150. 563.

Watson, Tamara L., Rachel A. Robbins and Catherine T. Best. 2014. Infant perceptual development for faces and spoken words: An integrated approach. Developmental Psychobiology 56. 1454–1481.

Weikum, Whitney. M., Tim F. Oberlander, Takao K. Hensch and Janet F. Werker. 2012. Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech perception. Proceedings of the National Academy of Sciences 109 (Supplement 2). 17221–17227.

Werker, Janet F. and Takao K. Hensch. 2015. Critical periods in speech perception: New directions. Annual Review of Psychology 66. 173–196.

Werker, Janet F. and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development 7. 49–63.

Wilson, Benjamin, Heather Slater, Yukiko Kikuchi, Alice E. Milne, William D. Marslen-Wilson, Kenny Smith and Christopher I. Petkov. 2013. Auditory artificial grammar learning in macaque and marmoset monkeys. Journal of Neuroscience 33. 18825–18835.

Zhang, Tie-Yuan and Michael J. Meaney. 2010. Epigenetics and the environmental regulation of the genome and its function. Annual Review of Psychology 61. 439–466.