What Infants Know and What They have to Learn about Language

JACQUES MEHLER[*], MARINA NESPOR[**] and
MARCELA PEÑA[†]

[*]International School for Advanced Science, Trieste, Italy and Emeritus
CNRS, France. E-mail: mehler@sissa.it; [**]University of Milan, Bicocca and
Interdisciplinary Center, Accademia dei Lincei, Italy; [†]Universidad de Chile and
International School for Advanced Science, Trieste, Italy.

The study of language acquisition during the first year of life is reviewed. We identified three areas that have contributed to our understanding of how the infant copes with linguistic signals to attain the most basic properties of its native language. Distributional properties present in the incoming utterances may allow infants to extract word candidates in the speech stream as shown in the impoverished conditions of artificial grammar studies. This procedure is important because it would work well for most natural languages. We also highlight another important mechanism that allows infants to induce structure from very scarce data. In fact, humans tend to project structural conjectures after being presented with only a few utterances. Finally, we illustrate constraints on processing that derive from perceptual and memory functions that arose much earlier during the evolutionary history of the species. We conclude that all of these machanisms are important for the infants to gain access to its native language.

Introduction

Chomsky's *Syntactic Structure* (1957)[1] and Eric Lenneberg's *Biological Foundations of Language* (1967)[2] had an enormous influence on several generations of students who were taught that humans have special mind/brain dispositions allowing them, and only them, to acquire the language spoken in their milieu. This was a momentous change of perspective given that behaviourists had imposed for the first half of the 20th century their restrictive views of learning. For instance, Mowrer,[3] a behaviourist who was already informed of Chomsky's critique of Skinner's work, see Chomsky[4] still tried to show that cognitive processes could be explained within the behaviourist tradition without reverting to 'cognition psychology', the term he used to designate

the new cognitive sciences. Indeed, according to Mowrer, everything can be handled, quite nicely, by means of the concept of 'mediation', a notion that he links to 'words', a token that links stimuli and responses (Ref. 3, p. 65). Moreover, imitation and statistics were given a prominent role in Mowrer's learning theory.

On the basis of production abilities of some avian species, the hypothesis was put forward that the ability of learning a language is not unique to humans. Parrots, mynah birds and other species can make vocalizations that resemble human utterances sufficiently as to allow listeners to decode what their sounds mean. Animal psychologists pursued their hopes that not only avian species could imitate speech sounds, but also that monkeys and apes could learn language if provided with the proper exposure and learning schedules. However, hard as they tried, none of the higher apes, despite their increased cognitive abilities, learned a grammatical system comparable to those underlying human languages.

What is it that makes it possible only for the human brain/mind system to acquire natural languages with great facility during the first few years of life? This question, apparently so simple to answer, still remains one of the great mysteries in the cognitive neurosciences. In this article we don't even attempt to answer such a difficult and all-encompassing question. Rather, we review some important mechanisms that facilitate language acquisition in healthy infants. We will also review some recent discoveries that advance our understanding of how the very young infant brain operates during language acquisition. Finally, we discuss why so many languages exist at the present time, even though language arose in only one location of the African continent some 70,000 years ago.

Any theory of language acquisition mechanisms must take into consideration some remarkable properties that attentive adults will notice. For instance, infants learn the mother tongue through mere exposure to the milieu in which they are born, or in any case in which they spend their first years of life. In fact, acquisition proceeds at its own pace regardless of formal training or other forms of couching. Of course, some parents may have the impression that it is they who are successfully teaching their infant several words per day, but this is mostly an unfounded impression. In fact, language onset becomes apparent at roughly the same age with all infants, as is also to be expected, since, in a non-exceptional milieu, biological dispositions are expressed, with minor variability, around the same mean age. Thus, clearly, the putative pedagogical efforts of parents are doomed to failure when dispensed to very small infants. Moreover, language learning is most effective during a 'window of opportunity' that does not last long and that is reminiscent of the instinctive learning that Gould and Marler[5] described when referring to the acquisition of songs in avian species. In fact, prior to exposure, birds produce a partial set of songs compared with normal adult birds. When the environment becomes available, it triggers the richer repertoire if the bird has reached the necessary neural maturation. Gardner[6] have shown that even when birds are exposed to a random walk of normal songs they will

regularize it to the 'syntactic' forms of grammatical songs to which they have never been exposed. This happens after 180 days of life, or much earlier if the birds receive a testosterone injection, demonstrating how a species-specific vocalization can arise either through a biological trigger or from the interaction with the milieu. Any observing parent knows that you cannot 'inject' say Quechua or Chinese to an infant. However, human language acquisition may depend largely upon setting in motion biological processes comparable to those that ethologists have uncovered.

Marler's and other ethologists' studies helped to shake the entrenched belief that the kinds of learning attested are all varieties of classical learning by association. This obviously is not the case. Not all fish have the ritualized kind of courting that is attested in three-spined sticklebacks in response to a bright red spot. Nor are all animals subjected to the kinds of imprinting that determines the duck response to a moving object just after hatching (see Ref. 7 and Ref. 8 respectively). Cognitive scientists began to view with sympathy the notion that learning varies greatly as a function of what members of a given species have to acquire given their habitat. The naturalistic ethological conception of learning lent support to the developing rationalist cognitive science and, in particular, to the suggestion that humans, like other animals, are endowed with special kinds of learning mechanisms, which are at the service of a faculty that is species-specific. The crux of much of the rest of this article explores some of the mechanisms that mediate language acquisition. While we do not have any reason to believe that any of those mechanisms are specific to language, their interaction might well be.

The problem of language acquisition

As mentioned above, there are no pills or injections that the traveller could be given to provide her/him with sufficient knowledge of Chinese when travelling to China, of Finnish when travelling to Finland, and so forth. Despite the specific human endowment to acquire language, we experience growing difficulties to acquire languages as we grow in age. This simple observation has led cognitive scientists to extricate themselves from classical learning theories. Indeed, why is it that at an age when the ability to learn mathematics or spelling is still absent, children are so proficient in learning a second or even a third language? And why is it that a few years later, younger adults will learn as many topics as needed while in college but will display severe problems when they have to acquire a new language? Even the sarcastic comments of Voltaire do not clarify matters. In his *Micromegas*,[9] he imagines a Cartesian thinker saying: 'L'âme est un esprit pur, qui a reçu dans le ventre de sa mère toutes les idées méthaphysiques, et qui, en sortant de là est obligé d'aller à l'école, et d'apprendre tout de nouveau ce qu'elle a si bien su et qu'elle ne saura plus (The soul is a pure spirit which receives in its mother's womb all metaphysical ideas and which on issuing thence, is obliged to go to school as it were

and learn afresh all it knew so well, and will never know again). In fact, cognitive scientists moved away from the traditional association learning theories, and started favouring species-specific learning mechanisms, even though these do not remove or solve all the problems either, as Voltaire's astute and ironic comment suggests.

In this chapter our purpose is to present some language learning mechanisms, which are the object of very active experimental investigations. In fact, Chomsky's theoretical arguments persuaded many cognitive scientists to explore the notion that only the human mind comes equipped with a specialized Language Acquisition Device (LAD). Chomsky (1965,[10] 1980;[11] see also Rizzi 1986[12]) proposed a theoretical framework called the Principle and Parameters (P&P) approach, which postulates universal 'principles' that apply to all natural languages, and a set of binary 'parameters' that have to be set according to the grammar of the surrounding language. Principles are mandatory for all natural languages, whereas parameters constrain the possible format of binary properties that distinguish groups of natural languages. The P&P formulation helped cognitive science propose realistic scenarios to explain how humans overcome the difficulties of first language learning.

Chomsky recently wrote[13] that

> the [P&P] approach suggests a framework for understanding how essential unity might yield the appearance of the limitless diversity that was assumed not long ago for language (as for biological organisms generally)… the approach suggests that what emerged, fairly suddenly, was the generative procedure that provides the principles, and that diversity of language results from the fact that the principles do not determine the answers to all questions about language, but leave some questions as open parameters.

Further along Chomsky states that

> We would expect, then, that morphology and phonology – the linguistic processes that convert internal syntactic objects to the entities accessible to the sensorimotor system – might turn out to be quite intricate, varied, and subject to accidental historical events. Parametrization and diversity, then, would be mostly – maybe entirely – restricted to externalization. That is pretty much what we seem to find: a computational system efficiently generating expressions that provide the language of thought, and complex and highly varied modes of externalization, which, furthermore, are readily susceptible to historical change.

The LAD would allow infants to set the different grammatical parameters to the value of their language of exposure. The proposal had a great influence on the field of language acquisition. Over the last two decades, however, it has become apparent that even if the LAD is a very attractive formal description of language acquisition, it does not specify in detail what it is that counts as a parameter, or even how many parameters need to be set in order to acquire a language. Nor did this formal proposal clarify how parameters are actually set given exposure to the linguistic data the learner receives.

Generative theorists have pointed out that the linguistic environment is far too impoverished for infants to learn the grammar of the native language on the basis of classical learning mechanisms. In fact, classical learning theories do not postulate a LAD or a specialized mechanism that would explain why it is that only humans can acquire grammatical systems. Moreover, to explain how infants acquire the grammar of their language of exposure, given that today over 6000 languages are spoken in different parts of the world, it is necessary to elucidate why it is so easy for infants to learn with equal facility any one of the languages that happens to be spoken in its surroundings. Biological observation suggests that the human species is endowed with special mechanisms to learn language. We thus need to understand how these putative mechanisms are deployed, how they interact with one another, and the order in which they are engaged. In other words, the experimental study of early acquisition has to be recognized as an essential area to judge theories for their formal pertinence and for biological validity.

Experimental studies show that considerable learning occurs during the first months of life (see, among others, Refs 14–18). Moreover, a number of investigators began to defend the notion of phonological and prosodic bootstrapping, that is, the notion that infants might explore the cues in speech that correlate with abstract properties of grammar (see, among others, Refs 19–23). The multiple cues that transpired from these studies make it possible to understand how and why it is that the notion expressed in the P&P approach might receive empirical support. In brief, research with very young infants supports the view that the signals received during the first year of life are potentially very rich in information and that the human brain, endowed as it is with mechanisms to extract the information carried by speech, will use them to discover abstract grammatical properties.

Below, we lay out in some detail some of the specific mechanisms that are used during the first years of life. First, we focus upon the faculty of the infant's brain to compute distributional properties of speech signals. Next, we explore the ability of the human mind to project generalizations on the basis of very sparse data. Finally, we explore some of the constraints that underlie both of the above mechanisms and how they handshake when they operate simultaneously.

Distributional computations

It has long been recognized that speech utterances tend to display statistical dependencies between phonological or lexical categories; some of those dependencies are verified in all spoken languages while others are language-specific.[24,25] Hayes and Clark[26] demonstrated that, in languages with multi-syllabic words, the transition probability (TPs) between word internal syllables tends to be higher than the TPs between the last syllable of any word and the first syllable of the next one. At first, this demonstration failed to attract much attention.

Cognitive scientists seemed to realize its potential importance only after Saffran *et al.*'s work[27] was published. Indeed, the results of Saffran *et al.* fulfilled the long-standing hope of discovering an unlearned neural computation that could provide insight into the ways in which very young infants manage to break into the constituents of their mother tongue. Consider how infants learn words from the speech signals in their environment. Most of the utterances are continuous. It is sufficient to look at the waveform of the utterances to be convinced that there are hardly any pauses, and that the few that there might be do not necessarily signal words. As adults we almost never find ourselves in a situation in which we have difficulties parsing utterances we hear – of course, utterances that belong to our native language – and thus we tend to think that infants also have an easy time parsing the utterances they hear. This is, however, not the case. Before babies are able to recognize a few words, they need to listen to utterances from the language spoken in their milieu, for almost a year. How then, do infants segregate the words that are contained in the utterances to which they are exposed? One possible suggestion is to claim that some words are presented in isolation. This, however, is a rare phenomenon.

Saffran *et al.*[27] discovered that eight-month-old infants can segment artificial continuous streams of syllables assembled in such a way that there are tri-syllabic 'words' that have high TPs between adjacent syllables while between words the TPs are lower. In the experiment, four 'words' were used; each word contained syllables of equal duration and intensity. Infants were familiarized with a stream for two minutes, and then were tested to establish whether they had extracted the 'words' on the basis of TPs. If they had, they should react differently to 'words' and to 'part-words', which contain the two last syllables of a word and the first syllable of another word, or the last syllable of a word and the first two syllables of another word. The results show that infants do react differently to 'part-words' and to 'words' in a head-turning paradigm.

Saffran *et al.*[27] convincingly demonstrated, then, that TPs, in the absence of any other cues, make it possible to parse such speech streams. We shall argue below that distributional cues are a very powerful kind of computation that can be exploited on line for segmentation as well for the extraction of other properties of the speech signal. In a recent publication, Gervain *et al.*[28] studied how infants could gain an understanding of the word order properties of their mother tongue before they had access to lexical items. The authors show that eight-months-old Japanese and Italian infants have opposite 'word-order' preferences after having been exposed to an artificial grammar continuous string composed of high frequency words and low frequency words in alternation, with each word being monosyllabic. After listening to the speech stream, Japanese and Italian infants were confronted with two-syllable items: one item had a high frequency syllable followed by a low frequency syllable; the other item had the two syllables in

opposite order. Japanese and Italian infants displayed a preference for different test items. This suggests that infants possess some representation of word order prelexically. The authors propose a frequency based bootstrapping mechanism to account for the above results. They argue that infants might build representations by tracking the order of function and content words, identified through their different frequency distributions.

Despite the power of distributional cues that are present in artificial speech, it is necessary to pursue such studies with methodologies that attempt to use strings that resemble more closely the utterances that the infants are processing in more realistic interactions. For instance, Shukla *et al.*[29] used an innovative design to have as much control of the stimuli as in an artificial grammar while adding prosodic cues to it. Shukla *et al.*[29] studied in greater detail how prosodic structures, and in particular intonational phrases (IP), interact with TP computations using an artificial speech stream. The IPs were modelled from natural Italian IPs. They found that when a 'statistical word' appears inside an IP, participants recall it without problems. This is not observed when one syllable of the same 'statistical word' appears at the end of an IP while the other two syllables appear in the next IP. Moreover, it is unimportant whether the Italian IPs are substituted by Japanese IPs, showing that the primacy of prosody over TPs is not necessarily the result of experience with one's native language.

Even while recognizing that the brain of very young infants has the capacity to track distributional properties present in signals, we still need to question whether distributional computations are sufficient to do away with much of the Chomskyan generative grammar notions, as was claimed by Bates and Elman.[30] Indeed, each infant has to learn her/his native language; the human brain must have the power to acquire the knowledge to produce and to understand any grammatical utterance of the language of exposure, a learning ability that no other animal brain displays. Even though animals have been shown to compute TPs, much as human infants do, they do not learn language.[31] Hence, although we do not question the importance of statistics, we ask which other mechanisms are necessary to attain grammar and how do they interact with one another.

Before we move to the next section we want to remind the reader that in natural languages non-adjacent dependencies hold between words as well as between morphological components. In the next section, we present experimental work showing that participants use non-adjacent TPs to parse continuous speech streams; we also present evidence that under some stimulating conditions participants discover and generalize to novel instances on the basis of regularities that were present during the familiarization phase. We will argue that the ability to project conjectures is another mechanism that is used in all kinds of situations and, in particular for extracting some grammatical properties, is a most singular accomplishment of the human mind.

Rules and generalizations

Peña *et al.*[32] set out to explore experimentally how the two mechanisms mentioned above (namely statistical computations and the projection of conjectures) interact with one another. In their first experiment the authors presented a continuous Saffran *et al.*-like stream to participants. The stream had high TPs between the first A syllable and the last C syllable of $A \times C$ like words, while the x varied. Participants were able to use the non-adjacent TPs to segment the stream into the $A \times C$ words, even though the xs of the test items had never occurred between an A and a C during familiarization. In their next experiment, the authors investigated whether participants could also extract the underlying regularity governing the 'words', namely that the first syllable A_i predicts the last one C_i regardless of the syllable(s) inserted between them. The first experiment evaluated the performance of the participants with a two-alternative forced choice test between a 'word' and a 'part-word' (a 'part word' can take two forms, namely *CAX* or *XCA*). The results demonstrated that participants chose words as being more familiar than part-words. In the second experiment, when participants were confronted with a 'part-word' versus a 'rule-word' (a 'rule-word' has the form AX*C where X* is a syllable that never appears in that position during familiarization) they failed to show any preference for one of the test items. Therefore, when continuous monotonous streams are used during familiarization, participants fail to generalize a regularity such as '*if A in first position then C in last position regardless of which syllable appears in between them*' although they can extract 'words' from the same streams relying on TPs between non-adjacent syllables.

Obviously, under more naturalistic situations, language learning infants never confront utterances composed of syllables of equal duration, equal intensity and with no pauses or modulations between the 'words'. Universally, speech utterances have complex prosodic structures. Although there are some prosodic variations between languages, what is never observed in well-formed utterances is the flat, isochronous syllables without pauses used in artificial grammar experiments. Consequently, we ought to try to establish how the brain operates when confronted with equally well-controlled stimuli when prosodic-like cues are inserted. For instance, how would participants behave if the continuity of the streams that Peña *et al.*[32] used in the above experiments is interrupted? Could a minor change in the stimulating conditions change the mechanisms the brain implements to extract information during the familiarization?

We hypothesized that when the brain is confronted with streams such as those used by Saffran and her colleagues, it will exploit the only mechanism that is left to try, namely computing distributional regularities to introduce, possibly, some structure into the otherwise homogeneous speech signal. Instead, when minimal pauses are introduced, participants may be able to learn that words start and end

with predictable syllables. As we will see, generalizations are usually projected on the basis of sparse data and not through purely statistical computations. Peña *et al.*[32] introduced silent pauses of 25 ms between the statistical 'words', which were tested in the two experiments described above. The very short pauses we inserted were not salient, and generally participants reported not hearing pauses when they were questioned at the end of the experiment. However, their processing mechanisms obviously were sensitive to the pauses, since the manipulation proved to be sufficient for participants to judge that 'rule-words' were far more familiar than 'part-words'. Even when the familiarization was reduced from 10 to 2 min, participants continued to judge 'rule-words' as more familiar, suggesting that mechanisms that verify regularities operate on the basis of very sparse data.

In an earlier experiment Marcus *et al.*[33] made a point that was, in part, close to one of Peña *et al.*'s points.[32] In fact, Marcus *et al.* were also persuaded that very young infants are capable of drawing generalizations on the basis of data short habituation. In fact, these authors proposed that to explain the acquisition of linguistic structures, next to simple distributional mechanisms, other mechanisms are indispensable. In particular, they argued that 7-month-olds use algebraic-like rules to extract and encode grammatical regularities. In order to support their claims, Marcus *et al.* carried out several experiments that were extremely influential because they suggested that 7-month-olds familiarized with structured items were able to generalize the structure to entirely novel items. For instance, in one experiment they familiarized the infants with *AAB* structured items (e.g. *fifigo*, *lalaru*, etc) and during the test phase the participants preferred *A*A*B** items (where the * indicates syllables that had never been used during the familiarization) over *A*B*B** items (*gofifi*, *rulala*, etc). The authors explain the infants' preference as arising because the participants notice that the *ABB* underlying structure differs from the *AAB* used during familiarization. Moreover, Marcus *et al.* proposed that statistical machines cannot account for the extraction and generalizations of structural regularities between some syllables, when the test concerns novel syllables. Even if different non-statistical mechanisms may be invoked, as we will see below,[34] it is clear that infants at some point in life begin making powerful generalizations. In fact, Marcus, like Chomsky[35] and Fodor[36] amongst others, holds the view that one of the most characteristic abilities of our species is to conjecture rules to learn regularities present in the input data we receive. Some of the authors cited suggest that rule-governed behaviour is at the basis of language acquisition.

Recently, Gervain *et al.*[37] explored the behaviour of newborn infants when confronted with items that share an *ABB* structure. They used a near-infrared spectroscopy device (NIRS) that allows observation of stimulus-related activation in 24 areas of the cortex, 12 over the left and 12 over the right perisylvian areas. The authors compared the blocks when infants listen to items whose first syllable differs from the second one, which reduplicates in third position (*ABB*),

with interleaved blocks during which the newborns listened to items whose three syllables differed from one another (*ABC*). The activations in response to the *ABB* items differ significantly from those the cortex of the newborns displays when confronted with the *ABC* items. The observed differences make it clear that the Marcus *ABB*-kind of regularity is processed very differently from the way an *ABC* pattern is processed. While the *ABB* items have a structural property that is easy to extract, the *ABC* items either do not give rise to a common structure (i.e. the regularity common to *barisu*, *fekumo*, etc), or it gives rise to many different structural properties, e.g. all items are tri-syllabic or all items have *only* different syllables, or all items have *only* CV syllables, etc. Newborns are not very apt at counting or having meta-linguistic abilities, nor do they formulate easily conjectures about what a series of items do *not* have in common. In contrast, the brain of the neonate, much like the nervous system of other animal species, has the ability to detect adjacent repetitions and to generalize from a series of items to novel ones.

Gervain *et al.*[37] also noticed that the newborns' cortical responses were different at the beginning compared with the end of the experiment. In fact, a time-course analysis indicated that the responses to the *ABB* items increase over time in some regions of interest, whereas response to the *ABC* items remains flat throughout the duration of the experiment. Taken together, these results suggest that newborn infants can extract the underlying regularity of the *ABB* items even though there are many different syllables that were used to assemble the individual items. Of course, there are, as we mentioned above, two possible mechanisms. The first postulates the computation of algebraic-like rules, such as 'all items have one syllable followed by a different one that duplicates'. The second is a mechanism to detect repetitions that, at least for the auditory modality, generalizes a structural property to items that repeat adjacently regardless of the token syllables that are used to implement the repetitions.

In a second experiment very similar to that described above, Gervain *et al.*[37] showed that the results were very different when the test items included only non-adjacent repetitions, i.e. items that conform to the *ABA* pattern, versus *ABC* items. The results indicate that there is no effect due to the extraction of an algebraic-like rule in this case. Nor is there either a hemispheric asymmetry effect or a change of activation pattern during the experiment. Indeed, when a time course analysis was carried out, the authors failed to find any effects. These results suggest that, at the initial state, a repetition detector fails to respond when repetitions are non-adjacent. This might signify that the repetition detection mechanism is not functional when there is a time-gap between the two *A* items that is greater than a few hundredth of a second. Further studies are being carried out to understand exactly the mechanism underlying the response to repetitions in very young infants.

Perceptual primitives, memory effects and other constraints

Over the last few years, an interesting discussion has arisen about the likelihood that the abilities of infants to acquire abstract knowledge are uniquely or mostly driven by algebraic-like rules. Obviously, focusing on the ability of using a mechanism of generalization that operates using propositional attitudes and algebraic-like rules seems to us to be an important move. It is unimaginable to hold the opposite view. How could science, the arts, politics and civilization emerge otherwise? Likewise, it seems most unlikely to view all acquisition of knowledge to be the result of just statistical learning of which associative learning is only one aspect. There is sufficient evidence showing that there is ample opportunity to extract and use such information during the course of language acquisition. However, language and knowledge acquisition are, in all likelihood, relying on several mechanisms that function either in parallel or during some time-locked phases of acquisition. Moreover, as we argued before, it appears that generalizations are drawn on the basis of very meagre information while the extraction of statistical information requires a greater amount of information. Likewise, there are other properties of the cognitive apparatus that influence the nature and the outcome of language acquisition. We call processing constraints those constraints that arise from the very nature of some perceptual and memory primitive processing operations.

It is clear that basic cognitive processes interact with language acquisition. It would thus be very surprising if such processes had failed to influence basic language acquisition processes. In fact, Endress *et al.*[34] demonstrate that only repetition-based structures with repetitions at the edges of sequences (e.g. *ABCDEFF* but not *ABCDDEF*) can be reliably generalized, although token repetitions can easily be discriminated at both sequence edges and sequence middles. These results suggest that there are interesting positional constraints that license spontaneous generalizations at certain positions of a speech stream but not at others. Why should that be? One answer to this question may reside in the fact that only the edges of a speech stream are salient, possibly an auditory 'gestalt-like' organization that has to be explained on a par with any other gestalt phenomenon. Otherwise, it is likely that only the edge positions of the stream count as true variables to which values can be assigned. The values of the other positions would refer to the initial and the final syllables.[38]

In several experiments, further support was gathered for such a view. For instance Endress and Mehler (in preparation) demonstrate that it is possible to use penta-syllabic items (A_iXYZE_i) to extract generalizations, namely to recognize that ($A_kX'Y'Z'E_j$) is accepted as a familiar item even it was never presented. This occurs only when each separate item is presented individually but not when a continuous stream is used. Moreover, when the second and fourth positions were used (as in XA_iYE_iZ) participants were *unable* to generalize when items were presented

individually. Nevertheless, they were able to use positions 2 and 4 to compute transition probabilities to segment a continuous stream. These results might not only clarify what participants do when confronted with artificial grammar learning experiments. They may also help explain why it is that most languages of the world use pre- and suf- fixes rather than infixes. Indeed, since edges of content words are more salient, it is reasonable to expect those positions to be privileged for inserting morphological markers. In fact, Endress and Bonatti[39] have described certain phenomena that only occur at the edges of items like those used in Peña *et al.*'s experiments.[32] Indeed, when participants are confronted with *AxC* items in non-continuous streams they tend to learn that there is a class of *A* items and another class of *C* items that can pair in ways that did not arise during familiarization, e.g. any A_i item can pair with any C_k item.

The above experiments ought to satisfy even the most critical readers of the existence of constraints that make edges very salient and highlight their role as variables, whereas the distance from the nearest edge is used to tag internal positions. It also shows that while TPs can be computed in any position, it is only in edges that generalizations can be projected. We suggest that a statistical machine cannot project generalizations and thus cannot explain these results.

Humans display a powerful perceptual mechanism derived from biological movement, namely, the iambic-trochaic law described in Nespor *et al.*[40] The signal determines grouping at the phrasal level and may thus be a cue to the value of the head-complement parameter: if phonological phrase stress is realized mainly by pitch and intensity (as in Turkish), then words are grouped trochaically, if mainly through duration and intensity (as in French) then they are grouped iambically.

Since in a pair of words of which one is the complement and one is the head, stress always falls on the complement, the particular physical manifestation of stress might indicate to the prelexical infant the relative order of head and complements. The same physical correlates of stress are found also intra-linguistically: in German, the phonological phrase stress of a complement pre-ceding its head is realized mainly through pitch and intensity, of a complement following its head mainly through duration.

Originally, the iambictrochaic law, proposed both for music[41,42] and for low levels of linguistic structure,[43] was based on intensity and duration only. Bion *et al.*[44] conducted a series of experiments to verify the role of pitch in grouping: do adults as well as infants indeed create prominent initial chunks if prominence is realized through pitch, and in prominent final chunks, if prominence is realized through duration? Similar experiments were also conducted on visual images and on gestures. All confirmed the hypothesized grouping. This suggests that the iambic-trochaic law is a further constraint based on perception across modalities and derived from biological movement: high frequency and short duration characterize the onset of groups and low frequency and longer duration their end.

In all the above examples of the primitives and the constraints we individualized, it appears that basic cognitive processes, e.g. memory, perception, gestalt-like emergents and biological movement are at the origin of their existence. These specialized tools modify dramatically the standard computational process displayed in higher mental processes and may help understand how the human faculties take the form they display. For instance, natural languages as well as logic may be the consequence of the interactions of statistics, rule-guided behaviour, and the primitive tools described above.

Conclusions

In this paper, we have reviewed a number of recently explored mechanisms that play an important role during language acquisition. It is clear that when we look at any one natural language, its structure is sufficiently complex that even the great grammarians, starting with Panini, and proceeding over Spinoza down to Jakobson, failed to understand the universal underlying regularities contained in grammar. Indeed, such regularities exist at the sound, lexical, syntactic, semantic and possibly pragmatical levels. Likewise, it is possible to describe distributional regularities at each one of these levels of description. The generative school of grammar, which Chomsky started in the late 1950s, was responsible for the discovery of many underlying regularities for many of the existing natural languages.

It is extraordinary that any neurologically undamaged infant early in life becomes as articulate as an adult if one compensates for the fact that young children do not have as sophisticated a vocabulary as do adults. It is possible that during development further acquisition is made, possibly by the growth of brain connectivity although, leaving aside issues of encyclopaedic knowledge, a child of 10 or 12 years of age is as proficient in language as any adult. Needless to say that no other animal comes equipped to acquire this uniquely human faculty.

We hope that the continuing exploration of the mechanisms deployed by human infants to learn the surrounding language, regardless of whether they are born blind or deaf, will yield a better understanding of how this becomes possible. At this time, however, there is a growing trend to bypass the synchronic questions to focus again on the evolution of language. In fact, Hauser *et al*.[45] were instrumental in rekindling the interest in how language evolved. In their paper they claimed that in order to make some sense of how language has evolved one has to differentiate the faculty of language in the broad sense (FLB) from the faculty of language in the narrow sense (FLN). The FLB includes the sensory-motor systems, the conceptual-intentional system and several other components, while the FLN, they conjecture, includes only recursion and this may turn out to be the only uniquely human component of language. In a more recent presentation already cited above, Chomsky[46] identifies linguistic recursion

with MERGE, one of the main components of his minimalist programme. Likewise, Pinker[47,48] has also argued for an evolutionary account of language, taking a very different perspective from that of Hauser *et al*. We welcome these different standpoints about how to conceive the study of the origins of language as well as why there are so many languages today. We believe, however, that the detailed experimental study of language acquisition and of how learning mechanisms that we might share with other species interact in a way unique to humans, complemented with the study of the neuroscience of development, is most likely to contribute concrete pieces of knowledge without which the grander project might not really succeed.

References

1. N. Chomsky (1957) *Syntactic Structure* (The Hague/Paris: Mouton).
2. E. Lenneberg (1967) *Biological Foundations of Language* (New York: John Wiley & Sons).
3. O. H. Mowrer (1960) *Learning Theory and the Symbolic Processes* (New York: Wiley).
4. N. Chomsky (1959) Review of Verbal Behavior by B. F. Skinner. *Language*, **35**, 26–58.
5. J. L. Gould and P. Marler (1987) Learning by instinct. *Scientific American*, **62**.
6. T. J. Gardner, F. Naef and F. Nottebohm (2005) Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science*, **308**(5724), 1046–1049.
7. N. Tinbergen (1951) *The Study of Instinct* (New York: Oxford University Press).
8. K. Lorenz (1958) The evolution of behavior. *Scientific American*, **119**, 67–78.
9. Voltaire (1979) *Romans et contes*, *Micromegas*, VII, pg 35, Gallimard, Ed Pleiade (Paris).
10. N. Chomsky (1965) *Aspects of the Theory of Syntax* (Cambridge, MA: MIT Press).
11. N. Chomsky (1980) *Rules and Representations* (New York: Columbia University Press).
12. L. Rizzi (1986) Null objects in Italian and the theory of pro. *Linguistic Inquiry*, **17**, 501–557.
13. N. Chomsky (2008) *The China Lectures* (In press).
14. P. K. Kuhl, K. A. Williams, F. Lacerda, K. N. Stevens and B. Lindblom (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, **255**(5044), 606–608.
15. J. F. Werker and R. C. Tees (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, **7**, 49–63.
16. P. W. Jusczyk (1997) *The Discovery of Spoken Language Recognition* (Cambridge, MA: MIT Press).

17. J. Mehler, P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini and C. Amiel-Tison (1988) A precursor of language acquisition in young infants. *Cognition*, **29**, 143–178.
18. J. L. Morgan (1996) A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language*, **35**, 666–688.
19. E. Wanner and L. R. Gleitman (1982) *Language Aquisition: the State of the Art* (Cambridge, UK: Cambridge University Press).
20. J. L. Morgan and K. Demuth (1996) Signal to syntax: an overview. In: J. L. Morgan, K. Demuth (eds) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 1–22.
21. M. Nespor (1995) Setting syntactic parameters at a prelexical stage. Proceedings of the XXV ABRALIN Conference, Salvador de Bahia.
22. M. Nespor, M. T. Guasti and A. Christophe (1996) Selecting word order: the Rhythmic Activation Principle. *Interfaces in Phonology* (U. Kleinhenz, Berlin: Akademie Verlag), pp. 1–26.
23. A. Cutler and J. Mehler (1993) The periodicity bias. *Journal of Phonetics*, **21**, 103–108.
24. G. K. Zipf (1936) *The Psycho-biology of language* (London: George Routledge).
25. G. A. Miller (1951) *Language and Communication* (New York: McGraw-Hill Book Company Inc).
26. J. R. Hayes and H. H. Clark (1970) Experiments on the segmentation of an artificial speech analogue. *Cognition and the Development of Language* (New York: Wiley), pp. 221–234.
27. J. R. Saffran, R. N. Aslin *et al.* (1996) Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.
28. J. Gervain, M. Nespor, R. Mazuka, R. Horie and J. Mehler (2008) Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. *Cognitive Psychology.*
29. M. Shukla, M. Nespor and J. Mehler (2007) An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, **54**(1), 1–32.
30. E. Bates and J. Elman (1996) Learning rediscovered. *Science*, **274**(13 December), 1849–1850.
31. M. D. Hauser, E. L. Newport and R. N. Aslin (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition*, **78**(3), B53–B64.
32. M. Peña, L. L. Bonatti, M. Nespor and J. Mehler (2002) Signal-driven computations in speech processing. *Science*, **298**(5593), 604–607.
33. G. F. Marcus, S. Vijayan, S. Bandi Rao and P. M. Vishton (1999) Rule-learning by seven-month-old infants. *Science*, **283**, 77–80.
34. A. D. Endress, B. J. Scholl and J. Mehler (2005) The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, **134**(3), 406–419.
35. N. Chomsky (1988) *Language and Problems of Knowledge* (Cambridge: MIT Press).

36. J. A. Fodor (1983) *The Modularity of Mind* (Cambridge, MA: MIT Press).
37. J. Gervain, F. Macagno, S. Cogoi, M. Peña and J. Mehler (In press). The neonate brain detects speech structure. *Proceedings of the National Academy of Science*.
38. N. Burgess and G. Hitch (2005) Computational models of working memory: putting long-term memory into context. *Trends in Cognitive Science*, **9**(11), 535–541.
39. A. D. Endress and L. L. Bonatti (2007) Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, **105**(2), 247–299.
40. M. Nespor, M. Shukla, R. van de Vijver, C. Avesani, H. Schraudolf and C. Donati (Under revision) Different phrasal prominence realization in VO and OV languages.
41. T. L. Bolton (1894) Rhythm. *American Journal of Psychology*, **6**, 145–238.
42. H. Woodrow (1951) Time perception. In: S. S. Stevens (ed.) *Handbook of Experimental Psychology* (New York: Wiley), pp. 1224–1236.
43. B. Hayes (1995) *Metrical Stress Theory* (Chicago and London: The University of Chicago Press).
44. R. H. Bion, M. Peña, A. Langus and M. Nespor (Submitted) Markers of prominence influence memory for auditory and visual sequences.
45. M. D. Hauser, N. Chomsky *et al.* (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science*, **298**(5598), 1569–1579.
46. N. Chomsky (1995) *The Minimalist Program* (Cambridge, MA: MIT Press).
47. S. Pinker (1997) *How the Mind Works* (Allen Lane: WW Norton).
48. S. Pinker and R. Jackendoff (2005) The faculty of language: what's special about it? *Cognition*, **95**(2), 201–236.

About the Authors

Jacques Mehler. After getting a PhD at Harvard University he became a member of the CNRS in Paris. In 2001 he became Professor of Cognitive Neuroscience at the International School for Advanced studies at Trieste where he heads the Language and Cognitive Development Laboratory. He has worked on language acquisition, specializing in the study of processing in neonates, using both behavioral and brain imaging methods.

Marina Nespor. After getting her PhD at UNC Chapel Hill she became Professor at the University of Amsterdam. In 1998 she became Professor at the University of Ferrara and in 2007 at the University of Milano-Bicocca. She has developed the area of sentence prosody and its consequences for language acquisition.

Marcela Peña Garay. After getting her medical degree from the University of Chile she specialized in pediatry. She got her PhD from EHESS in Paris. She is currently a Professor at the Center for Advanced Research in Education Universidad de Chile and at Pontificia Universidad Catolica de Chile, where she studies very early language development using brain imaging techniques.