ON DIFFERENT MECHANISMS INVOLVED IN THE ACQUISITION OF LANGUAGE

Marina Nespor[1], Jacques Mehler[2], Mohinish Shukla[3], Marcela Peña[2,4] and Judit Gervain[5]

[1]*University of Milano-Bicocca and Centro Linceo Beniamino Segre,*
[2]*SISSA – Trieste,* [3] *University of Rochester*
[4] *Universitad Católica de Chile,* [5]*University of British Columbia*

Abstract

In this paper, we discuss different mechanisms involved in language acquisition. We propose that both general learning mechanisms and language specific ones are involved in the acquisition of different aspects of language. While it is commonly assumed that the first type of mechanism is mainly engaged in the segmentation of connected speech and the identification of lexical items and the latter is mainly involved in the extraction of abstract regularities from the speech stream, we suggest that this division is not so sharp.

We will overview and discuss results in both domains. In addition, after reviewing different types of evidence concerning the asymmetry between consonants and vowels, we will discuss the identification of the segmental categories of the speech stream over which statistics are calculated in order to identify the boundaries of words. Given the specificity of these elements, the conclusion will be drawn that general learning mechanisms may be constrained by language specific representations.

1. Introduction

One of the achievements of the last thirty years in the field of language acquisition is the recognition of the importance of the first year of life for the understanding of the infant's steps towards the acquisition of the language(s) it is exposed to (among many others, Mehler and Dupoux 1990; Jusczyk 1997). Our research addresses the problem of the mechanisms that guide the infant during this process (Peña et al. 2002; Nespor et al. 2003; Bonatti et al. 2005).

To reach competence in the language they are exposed to, infants must acquire its lexicon and its grammar. Although these are by and large different tasks, both are complex, partly because, to newborns and young infants, a speech stream appears like a continuous flow of sounds, at least within intonation groups.

Once recognized that language acquisition starts long before infants utter their first words, cognitive scientists and developmental psychologists started investigating the cues in the speech stream that might allow infants to access the linguistic system they are exposed to, as well as the mechanisms involved in language acquisition.

Among the mechanisms known to play a role in the acquisition of the lexicon are general learning mechanisms such as statistical computations. A word internal pair of syllables is more frequent than a pair of syllables that belongs to two different words (Hayes and Clark 1970). Indeed, the segmentation of artificial continuous speech streams is achieved through the computation of the transitional probabilities between adjacent syllables (Saffran, Aslin and Newport 1996; Saffran, Newport and Aslin 1996). Once segmentation of the continuous speech stream is achieved, infants can start learning some words and construct the lexicon of their language of exposure. Imitation plays an important role in this process (Meltzoff and Prinz 2002). The computation of transitional probabilities is a domain general mechanism that operates also in the extraction of patterns of tones (Saffran, Johnson, Aslin and Newport 1999) or of visual stimuli (Fiser and Aslin 2001, 2002). We will show that even though this mechanism is exploited in the learning of different objects, in the case of language it is constrained by linguistic representations: transitional probabilities are computed over syllables and over consonants, but not over vowels (Bonatti et al 2005).

The acquisition of grammar presents very specific problems, given that children do not only repeat the sentences they hear, as is most often the case for words, but can produce sentences they have never heard before. Is the environment sufficiently rich as to inform the infant about the grammatical properties of its language in an unambiguous way? Chomsky's answer to this question is negative and he uses this argument, known as the *poverty of the stimulus* argument, to posit Universal Grammar (UG). UG bridges the gap between the insufficient stimulus and the fact that all humans without specific deficits acquire the grammatical complexities of their language of exposure in an effortless way. Our genetic endowment thus includes a language acquisition device specialized for language acquisition. The Principles and Parameters theory (P&P), first proposed by Chomsky (1981) conceives UG as a set of principles and a set of parameters. The first define the properties common to all the languages of the world, thus constraining the form human languages can take. The second are binary, mutually exclusive choices that define the dimensions along which languages vary. The basic idea is that small differences at an abstract level account for the large superficial variation observed in the different languages of the world. Given their universality, both principles and parameters are assumed to be part of the genetic endowment of the human species. Central to the acquisition of grammar is thus parameter setting, i.e. identifying, for each parameter, the value that describes a specific set of linguistic forms. The precise mechanisms that underlie this process, however, are left by and large unspecified in the generative literature.

Given that parameters define the different options a language might instantiate as to some major syntactic properties, in order for acquisition to take place, a trigger must be present in the language of exposure that allows the infant to set the parameter to the correct value. We are presently exploring the possibility that the acquisition of properties described through parameters is triggered by universal principles of perception rather than of grammar. That is, we are investigating the acoustic cues in the signal that could force the infant to learn either one or the other of the different grammatical properties described as parametric choices in the model of grammar.

Among the mechanisms proposed to be involved in the acquisition of grammar, there is also a mechanism that allows us to extract abstract patterns. Adults as well as infants appear to be able to extract algebraic-like rules to identify a pattern in a sequence, even though this pattern is physically instantiated in different surface forms. That is, both infants and adults are able to make generalizations on the basis of a pattern, even when it includes tokens that vary (Marcus, Vijayan, Rao and Vishton 2001; Peña, Bonatti, Nespor and Mehler 2002).

We believe that progress can be made in the understanding of language acquisition only by unifying the two different approaches mentioned above: while some aspects of linguistic competence are acquired through mechanisms specific to language, others are learned through general learning mechanisms, possibly constrained by linguistic representations. Of course, we will assume that if a learning mechanism is exploited in domains different from language, it needs not be postulated as belonging to the language acquisition device. In other words, an acquisition mechanism must be part of UG if and only if it concerns structures that cannot be learned through general learning mechanisms.

2. Segmenting the speech stream and constructing the lexicon

One of the important discoveries of recent years in the field of language acquisition concerns the ability of humans to segment the speech stream on the basis of statistical information (Hayes and Clark 1970). Participants of different experiments have been shown to be capable of inferring word boundaries in unfamiliar streams of monotonous syllables concatenated without intervening pauses, exclusively on the basis of statistical information (Saffran, Aslin and Newport 1996; Saffran Newport and Aslin 1996). The 'words' in these experiments are defined only on the basis of transition probabilities (TPs) between syllables. In all experiments, the 'words' are trisyllabic and all syllables consist of a consonant (C), followed by a vowel (V), that is, they instantiate the only syllable type that is present in all languages of the world. TPs are kept high between adjacent syllables that are word internal, and are low between two syllables that lie at the edges of two words.

The participants of these experiments are exposed for several minutes to continuous speech consisting of a sequence of 'words' delimited by dips in transitional probabilities and are then tested on their preference for words over partwords, where a partword is formed either by the last two syllables of a word and the first of the next or by the last syllable of a word and the first two of the next. The results show that word boundaries are inferred between syllables that do not always co-occur, but not between two syllables if the occurrence of the first faithfully predicts the occurrence of the second.

Experiments carried out with non-speech sequences, both acoustic (Saffran et al 1999) and visual (Fiser and Aslin 2001, 2002) in nature, also show that participants are able to identify portions of the sequences delimited by dips in TPs. All these experiments taken together indicate that the mechanisms used to chunk sequences to identify groups characterized exclusively by high TPs between their constituent elements belong to general learning and are thus not specific to the language faculty. For the segmentation of speech into words and the acquisition of the lexicon, general learning mechanisms thus appear to have an important role.

3. Why parameters?

One question we might ask is why parameters are needed to account for the way infants learn the grammar of their language of exposure. Why can infants not just learn the different properties of the grammar of the language they hear, rather than set parameters? The main reason for a parametric theory of language acquisition is that a single parameter does not in general describe one isolated property of a linguistic system, but rather a set of properties, often apparently unrelated on the surface. That is, if a single abstract grammatical property is present, several surface properties, not obviously deducible from one another, follow. In fact, if two or more properties are obviously linked, but it can be shown that infants can learn them one at a time directly from the signal and independently from the others, the need for parameter setting would have to be put into question; it would represent an unnecessary machinery, at least to account for how children acquire language. This would not imply that parameters are not adequate tools for a model of grammar or for language variation. It might however turn out that once more attention is paid to the actual mechanisms responsible for the learning of the variable properties of grammar, the function of parameters in language acquisition become superfluous.

To exemplify a variable property that has typically been given to illustrate a universal parameter, let us consider the head–complement parameter, which establishes whether heads precede or follow their complements. Related to this is the order of main and subordinate clauses. That is, a language expands either to the right or to the left both within phrases and within sentences. Mixed systems arise because of an on-going change (Kroch 2001 and various paper in van Kemenade and Vincent 1997; Sornicola, Poppe and Shisha-Hale 2000; Pintsuk, Tsoulas and Warner 2000). For the setting of this parameter, it has been proposed that a cue in the signal guides the infant, thus allowing a prelexical setting (Mazuka 1996; Nespor et al. 1996). In fact, if the parameter were set postlexically, e.g. once the child knows some nouns and some verbs, the setting mechanism would become superfluous (Mazuka 1996), if we assume that the X-bar theory is among the principles of grammar. Since X-bar states that all phrases have the same internal structure, if the object follows the verb, all complements follow their heads.

According to one proposal for the prelexical setting of the head–complement parameter (Nespor et al. 1996), the cue that would allow the infant to set it at the prelexical stage is the relative prominence within phonological phrases. Since head-complement languages have an iambic pattern at this level and complement-head languages a trochaic pattern (Nespor and Vogel 1986, 2008), hearing one of the two rhythmic alternations would allow the infant to know which type of language s/he is exposed to. Since 6-12 weeks old infants discriminate the two rhythms in delexicalized speech (Christophe et al. 2003), it is feasible that they could use this signal to set the parameter before knowing any lexical item and before being able to segment the speech stream. Under this view, in fact, each type of rhythm would be necessarily connected with a certain word order in universal grammar.

If, however, the location of prominence in phrasal rhythm is connected to word order because of general mechanisms, let us say of perception, than the need for the setting of a parameter disappears. This is what we would like to tentatively propose here. Our reasoning is the following. We know that humans group identical sounds that differ only in length into iambs, and identical sounds that differ only in intensity or pitch into trochees (Bolton 1894; Woodrow 1951; Allen 1975; Cooper 1976; Bell 1977; Jusczyk and Krumhansl 1993). This is known as the iambic-trochaic law (Hayes 1995). We also know that universally, the element containing new information is more prominent than elements that are given. In a sequence of head and complement in whatever order, it is most often the complement that bears new information (or the whole phrase, since focus projects (Nespor and Guasti 2002)). By hearing a sequence high–low, one would thus be forced to conclude that it represents a complement-head order; by hearing a sequence short-long, one would be forced to conclude that the sequence is head-complement (Nespor et al. submitted).

If things were as we conjecture, then there would be no need to have a parameter in UG, that is, two mutually exclusive possibilities in our *grammatical* genetic endowment. Rather, our perception mechanisms would allow us to identify the correct order of heads and complements, given the iambic-trochaic law and the fact that new information is more prominent than given information.

The conclusion that in the learning of major aspects of syntax, mechanisms not specific to the language faculty play an important role has yet to be proven (but see Gervain, Mehler, Nespor and Shukla, 2008). If this will turn out to be the case, the division of labor between general learning mechanisms and language specific ones will have to change its present configuration – specific for grammar vs. general for the lexicon – and become more articulated.

4. Extraction of abstract regularities

Parameter setting has never been proposed to account for the learning of all the regularities that emerge from grammatical structure. There are, for example, many morphological and morphosyntactic processes that must be acquired in order to have full competence of word formation, but that are not the consequence of parametric variation. Rather, they must be learned by directly inspecting the surface regularities and, on the basis of them, drawing generalizations.

Peña et al (2002) propose that whether participants in the experiments engage in one or the other of the two tasks can be influenced by subtle subliminal cues in the speech signal. It is shown that after a short familiarization to a continuous speech stream, adult listeners are able to segment it and isolate words, by computing TPs However, they fail to extract the structural regularities included in the stream either with the same or with a much longer familiarization. With the introduction of 25 msec pauses in between the words, these regularities are rapidly captured. These results are interpreted as showing that when subtle segmentation cues, possibly suggestive of prosody, are present, participants engage in the discovery of abstract, grammatical like regularities.

The discovery of the words that compose a stream and the discovery of structural regularities require different sorts of computations. The first is mostly statistical in nature and similar to that used to discover components in nonlinguistic domains. The second, used to discover grammatical-like regularities, is mostly rule based and possibly specific to language acquisition.

5. Consonants and the lexicon

In a recent paper, Nespor, Peña and Mehler (2003) propose that consonants and vowels have distinct functional roles in language. Consonants, through their rich qualitative distinctions, mainly identify lexical elements; vowels, through their rich quantitative distinctions, mainly signal syntactic properties (both universal and subject to parametric variation). This hypothesis is based on evidence coming from different disciplines that, with different methodologies, investigate human language and it has consequences for the elements on which it is conceivable that statistics are computed to segment the speech stream and consequently build the lexicon.

Work in linguistics has shown that consonants (Cs) and vowels (Vs) differ in their phonological representation (Clements, 1983; Goldsmith, 1995), their articulatory properties (Gafos 1999) and their acoustic properties (Stevens 2000). Separate tiers for Cs and Vs have been proposed in autosegmental phonology to account for the fact that certain phenomena make reference to one of the two categories ignoring the other. Given that adjacency is a general requirement on the description of most linguistic phenomena (among many others, McCarthy 1982; Goldsmith 1976; van Riemsdijk 1998), once there are separate representations for the consonantal and vocalic segments, these are adjacent on their respective tier and phenomena that refer to them are expected to exist. In particular, the consonantal tier has been proposed to account for the properties of lexical items in Semitic languages, where roots are consonantal and the intervening vowels constitute the morphological melodies that derive different words and word forms (McCarthy 1982). For instance, in Hebrew, the root *'gdl'* conveys the meaning of *'big'* but it is not the word *big*. It is just the root and the vowels that separate the consonants allow the generation of different word forms, as illustrated in (1).

(1) a. gadol ‘big’ (masculine adjective)
b. gdola ‘big (feminine adjective)
c. giddel ‘(he) grew’ (transitive verb)
d. gadal ‘(he) grew’ (intransitive verb)
e. higdil ‘(he) magnified’ (transitive verb)
f. magdelet ‘magnifier’ (lens)

Consonants and vowels also differ in their distinctive power. Cs are more distinctive than Vs because they outnumber them in the vast majority of linguistic systems, 20 Cs and 5 Vs being the most represented proportion. While vowels harmonize in many languages, as we will see below, Cs disharmonize rather than harmonize, thus increasing their distinctive power. In many languages there are phonological phenomena that apply in order to avoid sameness of features within words and these always concern consonants (Itô and Mester 1986; Yip 1988).

The fact that consonants carry most of the distinctive power in a lexicon and that a system like that of Semitic languages exists, make it conceivable that listeners are able to calculate transition probabilities on consonants, as they do on syllables. As we will see below, this is indeed the case: the general learning mechanism that allows humans to make use of statistical computations to extract words from fluent speech is also used on the consonantal tier (Newport and Aslin 2004; Bonatti et al 2005).

6. Vowels and grammar

Vowels, compared to consonants, have a lesser role in distinguishing lexical items for different reasons. Vs are more restricted in number than Cs in most systems. In many languages, not all vowels that occur in stressed positions may occur in unstressed positions as well. For example, in several Romance languages, a middle-low vowel in unstressed position is raised to high-low or a mid vowel to a high one if it looses word primary stress because of the addition of a suffix (Nespor 1993). Thus not only are Vs less numerous than Cs, but in certain positions they are even more restricted in occurrence. The extreme case is represented by systems, like the English one, in which all unstressed vowels are centralized to schwa. In addition, in many languages Vs harmonize throughout a domain, so that they have the function of demarcating a constituent, more than that of distinguishing word meanings. Demarcating a constituent, in fact, amounts to giving a cue to morphosyntactic structure.

Unlike the consonantal tier, mainly motivated to account for morphological patterns, as we saw above, the vocalic tier has indeed been proposed to account for phonological phenomena that spread over a phonological constituent, such as vowel harmony or tonal spreading (Goldsmith 1976), thus demarcating boundaries.

In addition, the prosodic properties of an utterance, mainly carried by vowels, might give a cue to syntactic structure (Selkirk 1984; Nespor and Vogel 1986, 2008). Thus two sentences with the same sequence of words but different syntactic structures can, in certain cases, be disambiguated by prosody. An example from Italian is given in (2) (from Nespor and Vogel 1986, 2008) and one from English in (3).

(2) La vecchia legge la regola

a. the old lady reads the law

b. the old law rules it

(3) The young bug flies

a. [The young bug] [flies]

The baby insect flies.

b. [The young] [bug flies]

Young people disturb flies.

Different parts of these utterances are stressed and lengthened, according to syntactic and consequently prosodic constituency. In general, in fact, the quantitative information carried by vowels through prosody gives a cue to syntactic structure in that prosodic constituents, though not necessarily isomorphic to syntactic constituents, do coincide with syntactic boundaries of different levels.

Prosodic boundaries are used in speech recognition (Salverda et al. 2003). The boundaries of phonological phrases are exploited on line in speech processing by both adults (Christophe et al. 2004) and infants (Gout et al. 2004). In addition it has been proposed that the relative prominence within phonological phrases (carried by vowels) gives a cue to the value of the head–complement parameter (Nespor and Vogel 1986, 2008).

Recently, it has also been proposed that the percentage of vocalic space occupied by vowels (%V) in the speech stream, in addition to indicating the rhythmic class and the richness of the syllabic repertoire (Ramus, Nespor and Mehler 1999), is also indicative of morphosyntactic properties: a high %V characterizes languages that have the object preceding the verb in the unmarked word order and that tend to have an agglutinative morphology (Shukla, Nespor and Mehler 2007).

Vocalic information thus does not only give cues to syntactic constituency, but may also help identifying some important characteristics of the language of exposure during the acquisition of phonology and morphosyntax. The fact that, as we will see below, TPs are not calculated on the vocalic tier in a way similar to that in which they are calculated on the consonantal tier, is one more indication that vowels do not play a major role in the acquisition of the lexicon.

7. Behavioral results concerning consonants and vowels

Infant's behavioral data provide evidence that consonants and vowels are processed in different ways during language acquisition. Throughout the first year of life, infants become specialized in the perception of the phonemes of their native language as compared to phonemes that are absent from it. Infants younger than 4 months discriminate consonants and vowels pertaining to their native language as well as consonants and vowels present in any other human language (Eimas et al. 1971; Kuhl et al. 1997; Morse et al. 1972). After 4 months, infants become particularly sensitive to the phonemes of the language they are exposed to, while they start reducing their capacity to discriminate phonemes absent from it (see Werker & Tees, 1999 for a review). However, this specialization occurs earlier for vowels than for consonants. From 6 to 8 months, the infants' capacity to discriminate foreign vowels is reduced (Polka 1994), while their capacity to discriminate foreign consonants starts decreasing only after 8 months of life (Werker 1984).

Behavioral data also show that normal adults exploit consonants and vowels differently in word recognition (Cutler 2000). When adults are allowed to change one phoneme to make a word from a non-word, they change a vowel more often than a consonant. For instance, presented with a non-word like *kebra*, listeners find it easier to recognize the word *cobra*, which results from the substitution of a vowel, than the word *zebra* that results from the substitution of a consonant. This behavior is similar for Spanish and for Dutch speakers, suggesting that consonants play a more distinctive role in word recognition, independently of the phonemic repertoire of the languages: in Spanish, in fact, there are many more consonants than vowels, while in Dutch there is a similar ratio between the two. These results suggest that vocalic information appears to constrain lexical selection less tightly (allows more potential candidates) than does consonantal information.

Newborns discriminate bisyllabic sequences as a function of their environment: inside a word or straddling phonological phrase boundaries (Christophe, Dupoux, Bertoncini and Mehler 1994; Christophe, Mehler and Sebastian Gallés 2001). Thus they might use prosodic information, mainly carried by vowels, as one of the cues to segmentation.

Phonological phrase prominence has been hypothesized to aid infants in the setting of the head–complement parameter (Nespor, Guasti and Christophe 1996) and indeed 6-12 weeks old infants discriminate utterances of the two languages crucially differing only in this prominence: syllabic structure and word stress are in fact similar and the sentences had been delexicalized so that no segmental information could distinguish the two sets of utterances.

8. Neuropsychological results

Using PET (positron emission topography), it has been shown that the increment of the cerebral blood flow in the left frontal operculum is larger when normal adults perform auditory detection tasks that incorporate stop consonants than steady-state vowel discrimination (Fiez 1995). Cortical stimulation techniques in patients tested during brain surgery, with no previous phonemic deficit, has shown that certain areas of the cortex are involved in consonantal and not in vocalic discrimination (Boatman 1997; Boatman 1995). Syllabic discrimination tasks accomplished on the basis of the detection of a change in consonants is highly compromised after a cortical stimulation (electrical interference) over a very small area of the superior temporal gyrus in the LH. On the contrary, syllabic discrimination carried out by the detection of a change in vowels, is less affected by the cortical stimulation in the same area and other brain regions in the peri-sylvian area.

Speech production studies in brain-damaged patients also suggest that categorical representation of vowels and consonants relies on different networks of the brain. A two-case brain damaged patient study (Caramazza et al. 2000) reports that the production of vowels and consonants is differently impaired independently of their distance in their acoustic properties, measured in terms of sonority. One patient with a lesion in the left parietal and temporal lobes and a small lesion in the right parietal lobe, made significantly more errors in the production of vowels than of consonants. The other patient with a lesion in the left supramarginal, angular and superior temporal gyri made significantly more errors in the production of consonants than in the production of vowels. The errors were related to phonemic categories and not to the sonority level of the phonemes.

9. Linguistic representations constrain general learning mechanisms.

Are general learning mechanisms constrained by representations specific to language? The empirical data mentioned above suggests that consonants and vowels have a different phonological behavior, play a different role in the interpretation of linguistic material – lexical and grammatical, respectively –, have a differential role in speech processing and are differently processed in the brain. In particular, consonants would preferentially contribute to the lexical processing of words while grammatical variations would rest mostly on the vocalic properties (Nespor, Peña and Mehler 2003; Bonatti, Peña, Nespor and Mehler 2005).

As shown before, research in the last few years has shown that a statistical mechanism is responsible for segmentation and construction of the lexicon: transition probabilities (TPs) between syllables allow subjects to isolate words (Saffran, Aslin and Newport 1996; Saffran Newport and Aslin 1996). This is a general learning mechanism, used also for musical and visual stimuli (Fiser and Aslin 2001, 2002). It has also been shown that TPs can be calculated exclusively on the consonants of speech streams consisting of CV syllables (Newport and Aslin 2004; Bonatti et al. 2005). This is not surprising, since the calculation of TPs on consonants would allow the extraction of words in Semitic languages. Newport and Aslin (2004) propose that TPs are calculated both exclusively on the consonants and exclusively on the vowels of streams of CV syllables. This result is expected, according to the authors, since there are languages in which phonological phenomena, such as vowel harmony involve only vowels.

We reasoned that since consonants, because of their qualitative distinctions, are specialized for lexical acquisition, TPs should be calculated on them. TPs should not, instead, be calculated on vowels, since they are very variable because of prosody and different phonological phenomena that obliterate their distinctions. Because of their main function in demarcating domains, rather than distinguishing lexical items, calculating TPs on their qualitative distinctions would serve no function in the acquisition of language. In addition, no languages have word roots formed only by vowels, a case that would be the counterpart of the consonantal roots in Semitic languages.

We carried out parallel experiments on consonants (with the intervening vowels varying) and on vowels (with the intervening consonants varying) and found that only in the first case can subjects segment and extract words (Bonatti et al. 2005).[1] On the basis of these experiments we conclude that even general learning mechanisms, such as the calculation of statistical probabilities, are constrained by linguistic representations: the consonantal, but not the vocalic tier is a representation on which TPs are calculated. The different results of Newport and Aslin (2004) on the one hand and Bonatti et al. (2005), on the other hand, are due, in our opinion, to the different material used in the two studies: while our material contains no adjacent repetition of two ‘words’, theirs does, It is thus the detection of repetitions rather than the calculation of TPs that might account for their results.

Recently we also asked if the perception of prosodic constituents like the intonational phrase (IP) might be used to constrain TPs. IPs account for natural break points in speech (Nespor & Vogel 1986, 2008). Since words are invariably aligned to the edges of IPs, we reasoned that subjects might use this information to constrain their search for words in a prosodically realized artificial speech stream. Indeed, we found evidence that indicates that high TP sequences that span IP boundaries are rejected as possible ‘words’ (Shukla, Nespor & Mehler, in press). These experiments provide further evidence that linguistic representations of the speech stream (in this case, prosodic

[1] In an experiment reported in Newport and Aslin (2004) participants extract words also on the basis of vowels. Their habituation streams, however, unlike ours, contain adjacent repetitions of the same word. Since with no repetition words are not extracted we interpret their results as being based on detections of repetitions.

domains) might constrain the process of segmentation using statistical cues.

10. Conclusions

In this paper, we reviewed our recent work on the different mechanisms involved in language acquisition, both the lexicon and grammar. We propose that, though general learning mechanisms such as the computation of transitional probabilities are exploited in the segmentation of continuous speech streams, the elements on which they are computed are restricted by representations that are specific to language. They are in fact computed both on syllables and on consonants, but not on vowels. This is in agreement with our proposal that each of the two different phonemic categories is preferentially involved in one aspect of language acquisition: consonants in the acquisition of the lexicon and vowels in the acquisition of grammar, most specifically syntax. Consonants, mainly alternating in quality, are more specialized for lexical distinctions. Vowels, because of being the main carriers of prosody, mainly alternating in quantity, mainly give cues to morphosyntactic properties.

We finally suggest that though mechanisms specific to language are involved in the acquisition of grammar, general learning mechanisms may also play a role. We thus draw the conclusion that the distinction between general mechanisms for the acquisition of the lexicon and language specific ones for the acquisition of grammar is less clear-cut then usually assumed.

Aknowledgements

The research presented in this article is supported in the framework of the European Science Foundation EUROCORES program *The Origin of Man, Language and Languages* and by the HFSP grant RGP 68/2002.

References

Allen, G. D. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics*, *3*, 75-86.

Bell, A. (1977). Accent placement and perception of prominence in rhythmic structures. *Southern California Occasional Papers in Linguistics*, *4*, 1-13

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, *6*, 145-238.

Bonatti, L., Peña, M., Nespor, M., & Mehler, J. (2005). The role of consonants and vowels in continuous speech processing. *Psychological Science*. 16. 6. 451-459.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein, & D. Lightfoot (Eds.), *Explanation in linguistics: The logical problem of language acquisition* London & New York: Longman.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *J Acoust Soc Am*, *95(3)*, 1570-80.

Christophe, A., Mehler, J., & Sebastian-Galles, N. (2001). Perception of prosody boundary correlates by newborn infants. *Infancy*, *2*, 385-94.

Christophe, A., Peperkamp, S., Pallier, C., Block, N., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*. 51. 523-547.

Cooper, W. E. (1976). Syntactic control of timing in Speech Production, Ph.D dissertation, psychology Department, MIT.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci*, *12(6)*, 499-504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn*, *28(3)*, 458-67.

Gafos, A. (1999). *The articulatory basis of locality in phonology*. New York: Garland Publishers.

Gervain, J., Mehler, J., Nespor, M., & Shukla, M. (in preparation). The psychological reality of parameter setting.

Gervain, J., M. Nespor, R. Mazuka, R. Horie and J. Mehler (2008) Bootstrapping word order in prelexical infants: A Japanese-Italian crosslinguistic study. *Cognitive Psychology.* 57. 56-74.

Goldsmith, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis*, *2*, 23-68.

Goldsmith, J. (1990). *Autosegmental and metrical phonology.* Cambridge: Blackwell.

Gout, A., Christophe, A., & Morgan, J. L. (2004) Phonological phrase boundaries constrain lexical access: Infant data. *Journal of Memory and Language*. 51. 548-567.

Hayes, B. (1995). *Metrical stress theory*. Chicago: Chicago University Press.

Hayes, J. R., & Clarke, H. H. (1970). Experiments on the segmentation of an artificial speech analogue. In J. R. Hayes (Ed.), *Cognition and the development of language.* New York: Wiley.

Itô, J., & Mester, A. (1986). The phonology of voicing in Japanese. *Linguistic Inquiry*, *17*, 49-73.

Jusczyk, P.W. & C. L. Krumhansl (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. Journal of Experimental Psychology: Human Perception and Performance. 19.3. 627-640.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge. MA. MIT Press.

Kemenade, A. van, & Vincent, N. (Eds.). (1997). *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy : Evidence of a domain general learning mechanism. *Cognition*. *83(2)*, 35-42.

Kroch, A. (2001). Syntactic change. In M. R. Baltin, & C. Collins (Eds.), *Handbook of contemporary syntactic theory* (pp. 629-739). London: Blackwell.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283(5398)*, 77-80.

Mazuka, R. (1996). Can a grammatical parameter be set before the first word? Prosodic contributions to early setting of a grammatical parameter. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition.* (pp. 313-30). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

McCarthy, J. (1982). *Formal problems in semitic phonology and morphology.* Doctoral dissertation, MIT. New York: Garland Press, 1985.

Mehler, J., & Dupoux, E. (1990). *Naître Humain*. Paris: Odile Jacob. English translation:

Mehler, J., & Dupoux, E. (1994). *What infants know*. Cambridge: Blackwell.

Meltzoff, A. N., & Prinz, W. (2002). *The imitative mind: Development, evolution and brain bases*. Cambridge: Cambridge University Press.

Nespor, M. (1993). *Fonologia*. Bologna: Il Mulino.

Nespor, M., Mehler, J., & Peña, M. (2003). On the different role of vowels and consonants in language processing and language acquisition. *Lingue e Linguaggio*, 221-47.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris. (2008) Berlin. Mouton. De Gruyter.

Nespor, M., M. Shukla, R. van de Vijver, C. Avesani, H. Schraudolf and C. Donati (submitted) Different phrasal prominence realization in VO and OV languages.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognit Psychol*, *48(2)*, 127-62.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298(5593)*, 604-07.

Pintzuk, S., Tsoulas, G., & Warner, A. (Eds.). (2000). *Diachronic syntax: Models and mechanisms*. Oxford: Oxford University Press.

Riemsdijk, H. v. (1981). On adjacency in phonology and syntax. (Ed.), *NELS XIII*

Riemsdijk, H. v. (1998). Head movement and adjacency. *Natural Language and Linguistic Theory*, *16*, 633-78.

Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, *17*, 501-57.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274(5294)*, 1926-28.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70(1)*, 27-52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606-21.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90(1)*, 51-89.

Shukla, M., M. Nespor and J. Mehler (2007) Interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*. 54.1. 1-32.

Shukla, M., Nespor, M., & Mehler, J. (in preparation). Grammar on a language map.

Sornicola, R., Poppe, E., & Shisha-Halevy, A. (Eds.). (2000). *Stability, variation and change of word-order patterns over time*. Amsterdam: John Benjamins.

Stevens, K. N. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Woodraw, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1234 - 1236). New York: Wiley.

Yip, M. (1988). The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry*, *19*, 65-100.