

SQUIB: ON PHONOLOGY AND PRELEXICAL MECHANISMS OF LANGUAGE ACQUISITION*

*Marina Nespor** (SISSA)****

Rhythm is order in movement

Plato, *The Laws*, Book II, 93.

Phonology is traditionally defined as the description of the grammatical competence of native speakers concerning the sound patterns of their language (Chomsky and Halle 1968). The field has subsequently been broadened to account for the abstract properties of phonology that are not specific to spoken language, but extend to the manual-visual modality in sign languages. A lot of attention has thus been devoted to the properties of the sensory-motor system of grammar - i.e. phonology - that are not specific to the auditory modality, that is, to spoken language, but extend to the visual modality. Abundant literature has explored this issue and showed that Sign Languages share many grammatical properties with spoken languages (among others, Brentari 1998; Nespor and Sandler 1999, Wilbur 1999; Sandler and Lillo-Martin 2006). This body of research has shown that parallel segmental and suprasegmental representations exist in the two linguistic modalities – the auditory and the visual.

Another issue largely discussed in more recent years, concerns the nature of the mechanisms involved in the acquisition of language. Which properties of language are perceived and acquired through mechanisms that are specific to language and which through mechanisms that are general, that is, used also in other domains? Arguments have been put forth both in favour of an innate language faculty (Mehler, Jusczyk, Lambertz, Bertocini, Amiel-Tison 1988), and in favour of general-purpose learning mechanisms (Elman, Bates, Johnson, Karmiloff-Smith, Parisi and Plunkett 1996, Tomasello 2000). No conclusions have been reached concerning this issue, i.e. that of the specificity vs. generality of the mechanisms that allow humans to perceive and acquire language. However, while the positions were rather extreme at the beginning of this dispute, most scientists would now agree that both types of mechanisms are involved (Nespor, Mehler, Shukla, Peña and Gervain 2009).

Within this line of research, I will first consider segmental representations and then suprasegmental, specifically, prosodic, ones.

* The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) / ERC grant agreement n° 269502 (PASCAL).

** nespor@sissa.it

*** Scuola Internazionale Superiore di Studi Avanzati, Trieste, Itália

Different levels – or tiers - of representation have been posited for consonants and vowels (Goldsmith 1990) because of evidence that there are phenomena that apply to vowels ignoring the intervening consonants and *viceversa*. More recently, it has been proposed that these two types of segments have different functions in the perception and acquisition of language. Consonants – quite stable in the speech stream - convey mainly lexical information, and vowels - the main carriers of prosody, thus very variable - mainly grammatical information (Nespor, Peña and Mehler 2003). An extreme case of the function of consonants (Cs) for the lexicon, and vowels (Vs) for grammar is constituted by Semitic languages, where exclusively Cs signal the meaning of roots, while the intervening Vs give grammatical information.

Though not limited to spoken language, since locations and movements in sign languages (SLs) have been compared to consonants (Cs) and vowels (Vs), respectively (Perlmutter 1992), to my knowledge, there is no acoustic correspondent of Cs and Vs in the non-linguistic world. The question, however, remains whether these elements specific to language are acquired through mechanisms that are themselves specific or through general ones.

One much investigated general mechanism of perception is the computation of transitional probabilities (TPs) between the adjacent elements of a string. The generality of this mechanism is proven by the fact that humans compute transition probabilities (TPs) on syllables (Saffran, Newport and Aslin 1996) as well as on non-linguistic tones (Endress 2010). The segregation of Cs and Vs has been demonstrated also on the basis of TPs: While TPs can be computed to segment words from continuous speech not only on syllables, but also on sequences of consonants, they cannot be computed on sequences of vowels (Bonatti, Peña, Nespor and Mehler 2005). A general perception mechanism can thus be constrained by representations – the segmental ones - that are purely linguistic in nature.

One way to prove the generality of perception mechanisms is to test non-human animals. In experiments with cotton-top tamarins, as participants, it has been shown that they compute TPs on syllables and on Vs, but not on Cs (Newport and Aslin 2004). Not being genetically programmed to learn a lexicon, tamarins compute TPs on Vs because they are more audible than Cs.

Adults, as well as infants, have also been shown to keep track of the relative frequency of functors and lexical items within a phrase and to group them according to the word order of their native language: functors first in Italian and functors last in Japanese (Gervain, Sebastián-Gallés, Díaz, Laka, Mazuka, Yamane, Nespor and Mehler 2013, Gervain, Nespor, Mazuka, Horie and Mehler 2008). Thus a general learning mechanism – keeping track of frequency – is suggested to aid infants in the acquisition of word order. In recent work, it has been shown that also rats can, to a certain extent, distinguish frequent from infrequent items, again proving the generality of the mechanism.

Prosody – at least rhythm - characterizes not only language in the two modalities – including gestures that accompany speech (Guellaï, Langus and Nespor, 2014) - but also non-linguistic events. There is rhythm in music as well as in many natural phenomena, ranging from the heartbeat to the waves of the sea, to the alternation of day and night and the way humans, as well as all animals, walk. But what is rhythm? To the best of my knowledge, the most general – thus best - definition of rhythm has been given by Plato: *rhythm is order in movement*. The task of scientists who investigate language

is to determine the elements that establish order in the movement of speech. Linguistic rhythm is hierarchical in nature. At the basic level it has been defined on the basis of the time occupied by vowels in the speech stream (%V) and of the standard deviation of consonantal intervals (ΔC) (Ramus, Nespors and Mehler 1999). On the basis of these measures it is possible to define the three traditional rhythmic classes: stress-timed (low %V and high ΔC), syllable-timed (higher %V and lower ΔC) and mora-timed (high %V and low ΔC). Infants have been shown to discriminate languages of different rhythmic classes, but not languages of the same rhythmic class. This capability can provide infants important information about the size of the syllabic repertoire of their language of exposure: large in stress-timed languages and progressively smaller in syllable-timed and mora-timed languages (Ramus et al. 1999).

At higher levels, linguistic rhythm is defined on the basis of the alternation of more and less prominent syllables. Human adults have been shown to group syllables alternating in pitch or intensity into trochees (strong – weak), and syllables alternating in duration into iambs (weak – strong), as predicted by the Iambic-Trochaic Law (ITL). The ITL has originally been proposed for the grouping of tones alternating in either intensity or duration in music (Bolton 1894, Woodrow 1951). This grouping is particularly important in language at the phrasal level for syntactic bootstrapping, since it offers cues word order. SOV languages have trochaic phonological phrase prominence mainly marked by pitch. SVO languages have iambic phonological phrase prominence mainly marked by duration. The type of prominence predicted by the ITL has been confirmed, at the level of the phonological phrase, for German, a language that, like Dutch, can have verbs both preceding objects and following them (Nespors, Shukla, van de Vijver, Avesani, Schraudolf and Donati 2008). The fact that different word orders are signalled differently also in languages in which both orders are found confirms that the ITL signals word order. In addition, it accounts for the fact that children exposed to one such language learn it at a similar pace observed in children exposed to a language that is uniformly either VO or OV.

A similar grouping has been shown not only for syllables (Bion, Benavides and Nespors 2011), but also for sequences of non-linguistic acoustic stimuli by participants of languages with different word orders (Hay and Dhiehl 2007; Langus, Seyed Allaei, Uysal, Primordian, Toro, Bion and Nespors, under review), as well as for sequences of non-linguistic visual stimuli (Peña, Bion and Nespors 2011). We also have recent experiments showing that duration and pitch at the phrasal level can be recognized from the analysis of the visual information provided by lips and mouth during speech production (Peña, Langus and Nespors, under review). Grouping governed by the ITL is thus neither restricted to language, since it applies also to non-linguistic tones, nor to the auditory domain, since it also governs the grouping of visual stimuli.

Non-human mammals, specifically rats, have also recently been tested for their sensitivity to the auditory ITL. Results show that they respect the ITL only for the grouping of sequences alternating in pitch, while they are at chance for elements alternating in duration (de la Mora, Nespors and Toro 2013). Since pitch is more audible than duration, these results remind us of those of tamarins with Vs and Cs briefly discussed above. Not having to rely on the ITL to bootstrap syntax, rats only group the most salient stimuli, that is, those alternating in pitch.

Prosody – at least rhythm – is thus present in many physical phenomena and it is - as segments - at least partly, perceived also by non-human mammals. The fact that rhythm gives a cue to word order and consonants to the lexicon is however, obviously, only specific to language.

Thus general mechanisms are applied to language in a way governed by linguistic representations: the statistical computation of TPs on consonants and syllables – but not on vowels - to learn the lexicon, and the rhythmic alternation defined by a general grouping principle - the ITL - to phrasal rhythm to learn one of the most basic properties of syntax that varies across languages: word order.

ACKNOWLEDGMENTS:

Thanks to Laura Bafile, Alan Langus and Jacques Mehler for discussions of different issues addressed in this paper.

REFERENCES

Bion, R.A.H, S. Benavides and M. Nespór. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, 54, 123-40.

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

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

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.

Chomsky, N. and M. Halle. (1968). *The Sound Pattern of English*. New York. Harper and Row.

de la Mora, D., M. Nespór and J.M. Toro. (2013). Humans and non human animals share the grouping principles of the Iambic – Trochaic Law. *Attention, Perception, & Psychophysics*. 75.1. 92-100.

Elman, J.L., E.A. Bates, M.H. Johnson, A. Karmiloff-Smith, D. Parisi and K. Plunkett. (1996). *Rethinking Innateness: a Connectionist perspective on development*. Cambridge, Mass. MIT Press.

Endress, A.D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, 135(2), 182-190.

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

Gervain, J., N. Sebastián-Gallés, B. Díaz, I. Laka, R. Mazuka, N. Yamane, M. Nespór and J. Mehler. (2013) Word frequency cues word order in adults: cross-linguistic evidence. *Frontiers in Language Sciences*, 4 (00689).

- Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Basil Blackwell.
- Guellaï, B., A. Langus and M. Nespôr. (2014). Prosody in the hands of the speaker.
- In I. Berent and S. Goldin-Meadow (eds.) Language by mouth and by hand. *Frontiers in Psychology*, Vol. 5.
- Hay, J.S.F. and R.L. Diehl. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception & Psychophysics*. 69 (1). 113-122.
- Langus, A., A. Saksida, D. Braida, R. Martucci, M. Sala and M. Nespôr. (2015). Spontaneous object and movement representations in 4-month-old human infants and albino Swiss mice. *Cognition*, Vol. 137, 63-71
- Langus, A., S. Seyed Allaei, E. Uysal, S. Primordian, J.M. Toro, R.A. Bion, and M. Nespôr (under review) What do we perceive natively?
- Mehler, J., P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini and C. Amiel-Tison. (1988). A precursor of language acquisition in young infants. *Cognition*. 29.2. 143-178.
- Nespôr, M. and W. Sandler. (1999). Prosody in Israeli Sign Language. *Language and Speech*. 143-176.
- Nespôr, M., M. Peña and J. Mehler. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*. 2. 221-247.
- Nespôr, M. and I. Vogel. (2007). *Prosodic Phonology*. Berlin. Mouton De Gruyter. Originally published in 1986 (Dordrecht. Foris).
- Nespôr, M., M. Shukla, R. van de Vijver, C. Avesani, H. Schraudolf and C. Donati. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*. 7.2. 1-28.
- Nespôr, M., J. Mehler, M. Shukla, M. Peña and J. Gervain. (2009). On different mechanisms involved in the acquisition of language. In M. Baltazani (ed.). *Proceedings of the 8th International Conference on Greek Linguistics*. University of Ioannina. http://www.linguist-uoι.gr/cd_web/arxiki_en.html
- Newport, E. L., M.D. Hauser, G. Spaepen and R.N. Aslin. (2004). Learning at a distance: II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49, 85-117.
- Newport, E.L., and R.N. Aslin. (2004). Learning at a distance: Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127–162.
- Peña, M., A. Langus and M. Nespôr (under review) I see what you say: syntax on your lips.

Peña, M. R.H. Bion and M. Nespors. (2011). How modality specific is the iambic-trochaic law? Evidence from vision. *Journal of Experimental Psychology: Language, Memory and Cognition*. 37.5. 1199-1208.

Perlmutter, D. M. (1992). Sonority and syllable structure in American Sign Language. *Linguist. Inquiry* 23, 407-442.

Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.

Saffran, J. R., E. L., Newport and R.N. Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Sandler, W. and D. Lillo-Martin. (2006). *Sign Language and Linguistic Universals*. Cambridge, UK: Cambridge University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* 74.3. 209-253.
Toro, J.M., M. Nespors, J. Gervain (under review) Frequency-based organization of speech sequences in a nonhuman animal.

Wilbur, R. B. (1999). Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech*, 42, 229-250.

Woodrow, H. (1951). Time perception. In S.S. Stevens (ed.) *Handbook of Experimental Psychology*. New York. Wiley. 1224-1236.

Recebido em: 14/04/2015

Aceito em: 24/04/2015