


Something Old, Something New: Combining Mechanisms During Language Acquisition

Current Directions in Psychological
Science
2016, Vol. 25(2) 130–134
© The Author(s) 2016
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0963721416629645
cdps.sagepub.com


Juan M. Toro

Catalan Institution for Research and Advanced Studies (ICREA), Pompeu Fabra University

Abstract

Infants face the paramount task of learning a language. Here, I review recent literature on two separate topics that suggest they use a combination of both evolutionarily old and new cognitive tools to face this task. Research on the principles that guide how humans and nonhuman animals group sequences of sounds has shown that we share with other species perceptual biases that we apply to linguistic stimuli. On the contrary, research on processing differences between consonants and vowels suggests humans, but not other animals, benefit from a “division of labor” across phonological representations. This division would help to extract regularities from the speech signal and facilitate language learning. The studies reviewed here provide support for the idea that perceptual biases together with language-specific representations guide the discovery of linguistic structures.

Keywords

speech perception, comparative cognition, perceptual bias, phonology

A recurrent question asked by researchers studying language acquisition is to what extent infants rely on highly specialized, uniquely human mechanisms to get a grip on language. The ability to create an infinite number of expressions by combining a relatively small number of elements, the hallmark of language, has been attested only in humans. It is therefore plausible that the cognitive tools supporting this ability are present in our species only. However, mounting evidence suggests that some of the abilities we use for language processing are present in other species. In the following pages, I will present recent findings from two different lines of research that contribute to our understanding of what is shared and what is uniquely human in language processing. The research described here has focused mainly on perceptual grouping principles and on differences across phonological representations. The results from several studies indicate that the process of acquiring a language involves both specialized constraints that are observed only in humans and general perceptual biases that are likely shared with other animals. I will thus briefly review studies with both humans and nonhuman animals to argue for a combination of evolutionarily old and new mechanisms during language acquisition.

Something Old

The first line of evidence comes from studies on perceptual grouping principles. When listening to an old clock, we tend to perceive the regular sequence of sounds it produces as a concatenation of *tick-tocks*. Similar perceptual grouping biases are captured by the iambic/trochaic law (ITL; see Fig. 1). The ITL was first proposed in the music domain (Bolton, 1894) but has also been applied to language (e.g., Nespor et al., 2008). The ITL states that sound sequences alternating in duration are grouped as iambs, with the strong element placed at the end of the sequence, thus producing a concatenation of short-long pairs. On the contrary, sequences alternating in intensity are grouped as trochees, with the strong element placed at the beginning of the sequence, producing pairs of high-low pairs, such as *tick-tock* in the clock. The way we perceive sequences of alternating sounds seems to conform to these principles.

Corresponding Author:

Juan M. Toro, Brain and Cognition Unit, Universitat Pompeu Fabra, C. Roc Boronat, 138 Barcelona 08005, Spain
E-mail: juanmanuel.toro@upf.edu

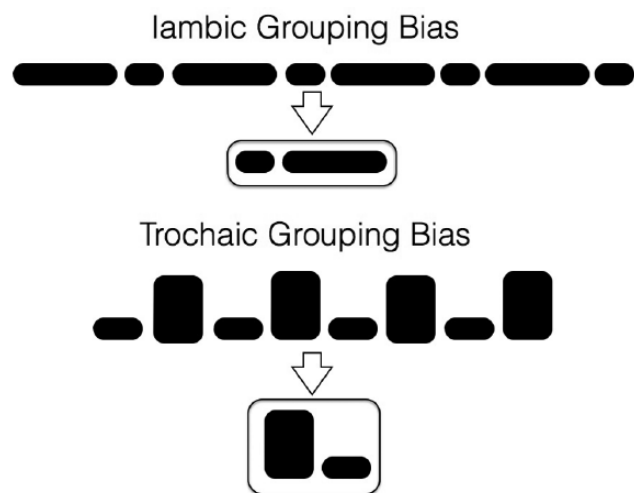


Fig. 1. Schematic representation of the grouping biases described by the iambic/trochaic law. Sequences of sounds varying in duration are grouped as iambs, with the short sound first and the long sound last. Sequences of sounds varying in intensity are grouped as trochees, with the more intense sound first and the less intense sound last.

Interest in the ITL in the context of language processing has grown in recent years. Researchers have found evidence that we tend to group sequences of syllables, tones, and visual figures following the principles described by the ITL (e.g. Bion, Benavides-Varela, & Nespors, 2011; Hay & Diehl, 2007; Iversen, Patel, & Ohgushi, 2008; Peña, Bion, & Nespors, 2011; Yoshida et al., 2010). Moreover, it has been hypothesized that such grouping biases could form the bedrock of how prosody is used to bootstrap syntactic structures (Nespors & Vogel, 2008). This is because acoustic realizations of linguistic prominence, such as higher intensity or longer duration, correlate with syntactic properties of a language. One example of such a correlation is the head-direction parameter: Languages in which verbs precede objects (e.g., English) are *head-initial*, whereas languages in which verbs follow objects (e.g., Japanese) are *head-final*, and differences in this syntactic parameter tend to correlate with how prominence is implemented acoustically (see Nespors et al., 2008). So, perceptual grouping principles such as the ones described by the ITL could be used to infer some syntactic properties of a listener's native language, such as the order in which words tend to be organized.

The fact that perceptual grouping principles apply across a wide array of stimuli, including syllables, tones, and visual figures, has led to the idea that they reflect general mechanisms that are shared across cognitive domains (e.g., Peña et al., 2011). If this is true, do the general grouping principles the ITL describes have their roots in mechanisms shared across species? In a series of experiments, de la Mora, Nespors, and Toro (2013) trained

animals to discriminate sequences of alternating strong and weak tones from sequences in which the same tones were randomly distributed. The animals were then tested on their preference for strong-weak or weak-strong tone pairs. In one experiment, the tones had a constant duration but their pitch changed, producing sequences of tones alternating between high and low pitch. During test, the animals responded more to strong-weak (high-low) pairs than to weak-strong (low-high) pairs. Thus, when the tones changed in pitch, the animals displayed a trochaic grouping bias mirroring that observed in humans. When a new group of animals was presented with sequences of tones with constant pitch but changes in duration (alternating short and long tones), the animals did not exhibit any grouping bias: They did not display any preference for weak-strong (short-long) or strong-weak (long-short) pairs.

Further studies on the ITL in humans led to the conclusion that prosodic properties of the listener's native language might modulate his or her sensitivity to certain cues. Whereas English and French adult speakers group sequences of tones varying in duration as short-long (displaying an iambic grouping bias; Hay & Diehl, 2007), Japanese speakers do not display any grouping bias (Iversen et al., 2008). This might be due to differences in phrasal structure across languages. In English, function words precede content words, so noun phrases have a structure analogous to iambic stress. In Japanese, on the contrary, function words are placed at the ends of sentences, resulting in a phrase structure analogous to trochaic stress (see Hay & Saffran, 2012; Iversen et al., 2008; Yoshida et al., 2010). So, exposure to a given set of language-specific prosodic features would constrain how sequences of sounds are grouped.

In this context, several lines of research suggest that the iambic and the trochaic biases have different developmental trajectories and are differently sensitive to experience. Whereas the trochaic bias has been observed early in life and across speakers of different languages (e.g., Bion et al., 2011), the iambic grouping bias appears later in life and is modulated by linguistic background (e.g., Bhatara, Boll-Avetisyan, Unger, Nazzi, & Höhle, 2013; Hay & Saffran, 2012; Iversen et al., 2008; Yoshida et al., 2010). Complementing these findings, recent experiments with animals have shown that it is only when given appropriate exposure to sounds alternating in duration that they display an iambic grouping bias (Toro & Nespors, 2015). Thus, the pattern of results observed in animals fits well with that observed in humans. They both suggest that specific experience is needed to develop the iambic grouping bias, while apparently no experience is needed to exhibit the trochaic bias. Data available so far from studies with human infants, adults, and other species suggest that by applying

the principles of the ITL to organize linguistic sounds, humans are taking advantage of a general perceptual mechanism that is present in other animals and has certainly not evolved for the purpose of language.

The ITL is just one of several cognitive capacities that humans use for language processing that do not seem to have evolved uniquely for linguistic purposes and are shared with other animals. Other examples include categorical perception (Kuhl & Miller, 1975), tracking statistical dependencies across elements in a sequence (Hauser, Newport, & Aslin, 2001), taking advantage of prosody to differentiate between speech sequences (Spierings & ten Cate, 2014), or using linguistic rhythm to discriminate among languages (Ramus, Hauser, Miller, Morris, & Mehler, 2000; for a review, see Yip, 2006). Even though advances have been made in this field, there is still much to be learned from comparative work on other species. For example, a promising line of research has explored parallels between birdsongs and speech (e.g., Doupe & Kuhl, 1999). The data available so far support the idea that human infants use some evolutionarily old mechanisms (present in animals distant from humans in the phylogenetic tree) to process the linguistic signal.

Something New

A different picture emerges from a second line of research exploring how phonology guides language learning. Consonants and vowels differ in several physical features. Vowels tend to be longer and carry more energy than consonants. But, more importantly, they seem to differ in the roles they play during language processing. When writing text messages we tend to omit vowels, but not consonants, in words (e.g., “txt mssgs”). We can r3pl4c3 v0w3ls 1n s3nt3nc3s and still be able to understand their meaning. Studies across different languages varying in their consonant:vowel ratio (e.g., English, Dutch, French, and Italian) suggest that this is not only a consequence of the fact that languages have more consonants than vowels, and that we rely more heavily on consonants than on vowels to identify words. Of course, this is a relative difference, as vowels can also be used to tell apart one word from another (we can tell that *ball*, *bell*, *bill*, *bull* are four different words though they differ only in their vowels). But experimental work has consistently shown that consonants have a heavier weight during lexical access than vowels. If participants are asked to create a word from a sequence of phonemes, they use the consonant frame (and not the vowel frame) as a reference (Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000). In visual masked-priming tasks, consonants in the prime word provide a more reliable source of information about the target word than vowels (New, Araujo, & Nazzi, 2008). This reliance on consonants for lexical identification is

even present early in life: Twenty-month-old infants tend to use consonants, instead of vowels, to learn new words (Nazzi, 2005), and 12-month olds rely heavily on consonant, and not on vowel, frames to identify words (Hochmann, Benavides-Varela, Nespor, & Mehler, 2011).

The relative weight that consonants and vowels have during lexical access seems to change during the first year of life. Five-month-old infants tend to rely more heavily on vowels than on consonants to identify their own names (Bouchon, Floccia, Fux, Adda-Decker, & Nazzi, 2015), and information carried by vowels is better recognized by newborns than information carried by consonants (Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012). However, by 12 months of age, infants have switched to focus on consonants to recognize words (e.g., Hochmann et al., 2011). Once the primary lexical role for consonants is established, our reliance on consonants to tell words apart contrasts with our difficulty to generalize abstract patterns over them. Both infants and adults find it easier to learn an abstract AAB rule (in which the two first elements in a sequence are repeated and the third is different) when applied over vowels than over consonants (Hochmann et al., 2011; Pons & Toro, 2010; Toro, Nespor, Mehler, & Bonatti, 2008). Thus, phonology seems to guide how linguistic information is extracted.

A way to explore the idea that phonological representations guide learning mechanisms is by assessing what happens when an organism putatively lacking such representations is presented with target stimuli. The question is, can we find a similar pattern of results in human and nonhuman animals? Or do the observed differences between consonants and vowels emerge exclusively from the way humans represent the linguistic signal? In a series of experiments, rats were trained to discriminate between consonant-vowel-consonant-vowel-consonant-vowel nonsense words in which either the consonants or the vowels followed an AAB rule. After training, they were tested with novel words that either followed the rule or did not. Animals correctly discriminated between the new words independently of whether the rule was implemented over consonants or over vowels (de la Mora & Toro, 2013). Human participants tested with exactly the same stimuli learned the rule as applied over the vowels, but not over the consonants. Animals thus outperformed humans in this rule-learning task when the rule applied to consonants. Interestingly, cotton-top tamarin monkeys also relied on different phonetic cues from humans while extracting regularities from a speech stream. Whereas humans preferentially compute statistics over consonants (Bonatti, Peña, Nespor, & Mehler, 2005), tamarin monkeys compute them over vowels (Newport, Hauser, Spaepen, & Aslin, 2004). The emerging picture is that, when present, phonological representations would

constrain how we extract information from the signal. Lacking such representations, animals would not benefit from these constraints.

What other cognitive tools involved in language processing are not likely to be found in nonhuman animals? There are plenty of good candidates, but most notably, it is the ability to combine elements at one level to create units at a different one (to combine phonemes to create syllables, syllables to create words, and words to create sentences; e.g., Hauser, Chomsky, & Fitch, 2002). However, this is a field of inquiry that is still growing, and much more research is needed to clearly establish the limits of what is uniquely human during language processing.

Putting the Pieces Together

Research on the basic mechanisms human infants use to learn the complexities of their native language has considerably advanced in the past years. The extent to which these mechanisms are present in other animals is still debated, and more data from different species is certainly necessary. Testing a wider array of species, including vocal and nonvocal learners, would help us to understand the evolutionary history of these traits. For example, it would be important to distinguish which mechanisms are shared across different species by common descent and which are shared across species but evolved separately (see Fitch, 2005). Together with a growing body of recent findings on comparative cognition and language development, the results reviewed above point toward the idea that infants use a diverse cognitive tool kit when approaching the task of acquiring a language (see also Endress, Nespor, & Mehler, 2009). Research on the grouping principles described by the ITL has shown that infants take advantage of perceptual biases that are shared across species, including animals distant from humans in the phylogenetic tree (research on songbirds would be very welcome in this area). When applied to linguistic stimuli, these principles would help infants to bootstrap syntactic parameters of their native language, such as word order (e.g., Nespor & Vogel, 2008). Studies on functional differences between consonants and vowels have suggested that humans, but not other animals, benefit from a “division of labor” provided by phonology. This division would help infants to focus on relevant sources of information. Lacking these language-specific representations, nonhuman animals process all phonetic categories as equivalent and do not benefit from their differences.

General perceptual biases together with language-specific representations seem to guide infants’ discovery of linguistic structures. Research on what is shared across species and what is uniquely human suggests that the

use of cognitive tools with varying degrees of specialization facilitates the paramount task of learning the complexities of human language.

Recommended Reading

- Hauser, M., Chomsky, N., & Fitch, T. (2002). (See References). An article that explores what aspects of the language faculty might be found in other species and puts forward the influential hypothesis that only recursion is uniquely human in language.
- Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). (See References). A study demonstrating different grouping patterns across speakers of English and Japanese.
- Nespor, M., & Vogel, I. (2008). (See References). A general and accessible book describing the correlation between prosody and syntactic structures.
- Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. (2008). (See References). An empirical work on the functional differences between consonants and vowels.
- Yip, M. (2006). (See References). A comprehensible review of several studies exploring the extent to which some phonological features might be based on mechanisms shared across species.

Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was funded by the European Research Council (ERC Starting Grant agreement number 312519).

References

- Benavides-Varela, S., Hochmann, J., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn’s brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences, USA*, 109, 17908–17913.
- Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T., & Höhle, B. (2013). Native language affects rhythmic grouping of speech. *Journal of the Acoustical Society of America*, 134, 3828–3843.
- Bion, R., Benavides-Varela, S., & Nespor, M. (2011). Acoustic markers of prominence influence infants’ and adults’ segmentation of speech sequences. *Language and Speech*, 54, 123–140.
- Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, 6, 145–238.
- Bonatti, L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16, 451–459.
- Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 month. *Developmental Science*, 18, 587–598.
- Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on

- lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, 28, 746–755.
- de la Mora, D., Nespors, M., & Toro, J. M. (2013). Do humans and non-human animals share the grouping principles of the Iambic-Trochaic Law? *Attention, Perception, & Psychophysics*, 75, 92–100.
- de la Mora, D., & Toro, J. M. (2013). Rule learning over consonants and vowels in a non-human animal. *Cognition*, 126, 307–312.
- Doupe, A., & Kuhl, P. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631.
- Endress, A., Nespors, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13, 348–353.
- Fitch, T. (2005). The evolution of language: A comparative review. *Biology & Philosophy*, 20, 193–230.
- Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Hauser, M., Newport, E., & Aslin, R. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B53–B64.
- Hay, J. F., & Diehl, R. (2007). Perception of rhythmic grouping: Testing the Iambic/Trochaic Law. *Perception & Psychophysics*, 69, 113–122.
- Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, 17, 610–641.
- Hochmann, J., Benavides-Varela, S., Nespors, M., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, 14, 1445–1458.
- Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, 124, 2263–2271.
- Kuhl, P., & Miller, J. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, 190, 69–72.
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, 98, 13–30.
- Nespors, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio*, 2, 1–29.
- Nespors, M., & Vogel, I. (2008). *Prosodic phonology*. Dordrecht, The Netherlands: Foris.
- New, B., Araujo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, 19, 1223–1227.
- Newport, E., Hauser, M., Spaepen, G., & Aslin, R. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a nonhuman primate. *Cognitive Psychology*, 49, 85–117.
- Peña, M., Bion, R., & Nespors, M. (2011). How modality specific is the Iambic-Trochaic Law? Evidence from vision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1199–1208.
- Pons, F., & Toro, J. M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition*, 116, 361–367.
- Ramus, F., Hauser, M., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349–351.
- Spierings, M., & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proceedings of the Royal Society B: Biological Sciences*, 281, Article 20140480. doi:10.1098/rspb.2014.0480
- Toro, J. M., & Nespors, M. (2015). Experience-dependent emergence of a grouping bias. *Biology Letters*, 11, Article 20150374. doi:10.1098/rsbl.2015.0374
- Toro, J. M., Nespors, M., Mehler, J., & Bonatti, L. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, 19, 137–144.
- Yip, M. (2006). The search for phonology in other species. *Trends in Cognitive Sciences*, 10, 442–446.
- Yoshida, K., Iversen, J., Patel, A., Mazuka, R., Nito, H., Gervain, J., & Werker, J. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition*, 115, 356–361.