ADDRESSING THE "VOCAL" IN VOCAL LEARNING

PEDRO TIAGO MARTINS[*1,2] and CEDRIC BOECKX[1,2,3]

[*]Corresponding Author: pt.martins@ub.edu
[1]Section of General Linguistics, University of Barcelona, Barcelona, Spain
[2]University of Barcelona Institute of Complex Systems, Barcelona, Spain
[3]Catalan Institute for Research and Advanced Studies, Barcelona, Spain

1. Vocal learning is usually about the learning

Vocal learning is hard to define as a phenotype. It is not known how pervasive it is in the animal kingdom, but it is generally considered to be present in at least three families of birds (songbirds, parrots, and hummingbirds), and in some mammals, which include cetaceans (Janik, 2014), pinnipeds (Ravignani et al., 2016), bats (Vernes, 2017) and elephants (Stoeger & Manger, 2014), and humans. This is understood to refer to (at least) vocal production learning, one of the 3 subtypes in what has become an well established typology, after (Janik & Slater, 2000):

vocal comprehension learning ability to associate a signal with a behavioral response

vocal usage learning ability to learn the context in which a vocalization can be used

vocal production learning ability to significantly modify vocalizations on the basis of experience

The crucial differences between these types of vocal learning have to do with *what is learned*. The first two types (*comprehension* and *usage*) pertain to the association of existing signals with new contexts, while the third kind pertains to the learning of the signals themselves. This could involve either a modificiation, as long as it can be shown that such a modification is learned from experience (e.g. auditory), or the production of novel vocalizations altogether. It is sometimes considered that only the production of novel calls can be considered to be vocal learning (Fitch, 2010). The questions of what counts as novel, what is learned, and what are the minimal units of animal calls for each species are not trivial, and they very often influence which vocal learning category a behavior will be ascribed to. Naturally, these questions inform much of the work on vocal learning.

2. The vocal in vocal learning

However, there is an important aspect of vocal learning that is rarely discussed or defined with little justification, which has to do not with the learning of signals but with their source. In other words, what "vocal" actually is. In a sense, this should not be important, if what is interesting about the phenotype is the learning. But in practice this issue goes beyond terminology: it rigidly contrains what different researchers count as a behavior to be studied, and has ramifications that go from phylogeny to the neurobiology and beyond. We identify different possibilities, which go from most limitative to most permissive:

- phonatory muscles: ability to produce sound by controling the larynx (e.g. humans) or syrinx (e.g. songbirds) (Elemans et al., 2015)
- any part of vocal tract: ability to produce sound by controling structures that could include the phonatory muscles but could be limited to the structures in the upper vocal tract, such as orangutan whistling (Lameira et al., 2013)
- any orofacial structure: ability to produce sound by using orofacial structures that might be outside of the vocal tract proper, such as the nasal tract in elephants (Stoeger & Manger, 2014)
- "artifical extension" of the vocal tract: ability to modify sounds by using external means outside of orofacial structures, such as hands or foreign objects, (e.g. orangutan (Hardus, Lameira, Schaik, & Wich, 2009))
- shortand for sound volitionally produced by animals, regardless of anatomy but usually assumed to involve any or all the means listed above

We contend that there are no *apriori* reasons for choosing one option over the other, and that each of them yield completely different approaches to vocal learning that have ramifications beyond behavior, and make classification and comparative work more challenging. One clear case is neurobiology: it is generally assumed that vocal learning involves a direct forebrain projection to the phonatory muscles (Fitch, 2010; Jarvis, 2007), and this idea has been used as a guideline for all species (Petkov & Jarvis, 2012). Species not shown to volitionally produce sounds using these anatomical and brain structures are often considered to be non-vocal learners, *regardless of their learning abilities pertaining to conspeficic calls*. (Ghazanfar, Liao, & Takahashi, 2019).

We discuss the implications of this state of affairs for the classifcation of different species in vocal learning, for the study of its neurobiology, and for devising evolutionary theories of this phenotype and other, closely related ones, with language as an important case.

References

Elemans, C., Rasmussen, J., Herbst, C., Düring, D., Zollinger, S., Brumm, H., Srivastava, K., Svane, N., Ding, M., Larsen, O., Sober, S., & Švec, J. (2015). Universal mechanisms of sound production and control in birds and mammals. *Nature Communications*, *6*(1), 8978.

Fitch, W. T. (2010). *The evolution of language.* Cambridge: Cambridge University Press. (OCLC: 703204991)

Ghazanfar, A. A., Liao, D. A., & Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Animal Behaviour*.

Hardus, M. E., Lameira, A. R., Schaik, C. P. V., & Wich, S. A. (2009). Tool use in wild orang-utans modifies sound production: A functionally deceptive innovation? *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1673), 3689-3694.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60-65.

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1–11.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, *148*(1), 35–44.

Lameira, A. R., Hardus, M. E., Kowalsky, B., de Vries, H., Spruijt, B. M., Sterck, E. H. M., Shumaker, R. W., & Wich, S. A. (2013). Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *The Journal of the Acoustical Society of America*, *134*(3), 2326-2335.

Petkov, C. I., & Jarvis, E. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*.

Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. S., & de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. *Frontiers in Neuroscience*, *10*.

Stoeger, A. S., & Manger, P. (2014). Vocal learning in elephants: Neural bases and adaptive context. *Current Opinion in Neurobiology*, *28*, 101-107.

Vernes, S. C. (2017). What bats have to say about speech and language. *Psychonomic Bulletin & Review*, *24*(1), 111-117.