Journal of Phonetics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Biology, culture, evolution and the cognitive nature of sound systems

Bart de Boer

AI-lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

ARTICLE INFO

Article history:
Received 4 September 2014
Received in revised form
27 May 2015
Accepted 2 July 2015

Keywords:
Speech sound systems
Evolution
Cultural evolution
Iterated learning
Computer models

ABSTRACT

This paper reviews the importance of the interaction between cultural evolution, biological evolution and individual cognition in understanding the cognitive nature of speech sound systems. Because of the effect of cultural evolution, typological properties of languages do not reflect individual cognitive mechanisms directly. In addition, the interaction between cultural evolution and biological evolution deeply influences what kind of cognitive adaptations to speech can evolve.

Theoretical work and computer simulation have shown that at least two kinds of adaptations to speech and language can evolve. One consists of weak biases to discrete features of language (such as word order) that convey a functional advantage. The other consists of stronger adaptations involving continuous traits in which language and biology can co-evolve (the vocal tract being a possible example of such a co-evolved adaptation). Experimental work is underway to identify how exactly cultural and biological evolution interact in human speech and language, and what cognitive mechanisms (if any) may have undergone selective pressure related to speech and language.

The paper reviews a number of studies that take the evolutionary perspective, focusing notably on agent-based computer simulations and on experimental work that simulates evolution in the laboratory or experimental work that investigates the interaction between individual learning behavior and cultural transmission directly. The paper argues that taking the evolutionary perspective (both cultural and biological, as well as their interaction) into account is necessary for a full understanding of the cognitive nature of language and speech.

1. Introduction

In order to understand the nature of speech sound systems we have to come to terms with the fact that they have a cognitive, a physical and a cultural component. These components have been shaped and have influenced each other through developmental and evolutionary processes (Christiansen & Kirby, 2003), see Fig. 1. Biological evolution shapes the brain and the physical apparatus to produce speech. Cultural evolution shapes the systems of speech sounds themselves as they are transferred through learning from speaker to speaker. In addition, culture and biology influence each other through co-evolution: biological adaptations determine which speech sounds are easy to produce, perceive and learn (and thus which are likely to be favored by cultural evolution) while the ability to communicate effectively and efficiently partly determines biological fitness. The interaction between cultural and biological evolution makes it difficult to disentangle which aspects of speech are due to cognitive mechanisms, which are due to cultural processes and which are due to physical properties. Moreover, it may be difficult to ascertain which mechanisms used in producing, perceiving or learning speech are specific to speech and which already existed before speech. This paper reviews work that seeks to answer these questions, as well as the computational and experimental methods used.

The cognitive and physical machinery are both determined by biology. Cognitive mechanisms have to do with mechanisms in the brain. These can be such mechanisms as the ones that allow us to sequence the rapid movements needed to control the vocal tract, the ones that allow us to split up streams of continuous speech into its building blocks (and to remember and reuse those building blocks), the ones that allow us to extract rules from the patterns that we observe and many more. Physical properties have to do with the production and perception of the physical signals. These can be our specialized vocal tract, our breathing apparatus and our hearing apparatus. The difference between cognitive and physical adaptations is not always clear-cut, as is illustrated by the proposed neural mechanisms that allow us to control our vocal folds voluntarily (see Fitch, 2010, ch. 9 for a review) or the neural

E-mail address: bart@ai.vub.ac.be

0095-4470/$ - see front matter © 2015 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2015.07.001

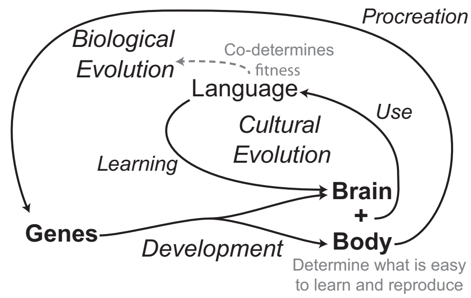

Fig. 1. The different processes and influences in evolution of language. Physical aspects are indicated in bold, while processes are indicated in italic. Language itself, being concrete but ephemeral is indicated in regular fonts. Arrows indicate influences and transmission of information.

mechanisms in the inner ear. This is of course to be expected; after all, cognitive adaptations are just physical adaptations to the brain and the nervous system.

It should be noted that biological mechanisms are not solely determined by genetics. They are shaped through a complex developmental process (which in itself is partly determined by genetics) in interaction with the environment – a process which is often abstracted away from in the work reviewed here. Developmental processes and interaction with the environment play a role in many biological systems, but the ones involved in language are unique in the sense that the environment that shapes them is partly determined by culturally transmitted information, which itself is partly determined by the physical and cognitive properties of humans.

Cultural processes are related to speech and language being bodies of knowledge that are shared and transmitted in a community of speakers (rather than in a single speaker's brain). Self-organization, cultural evolution (Steels, 1998) and historical contingency are examples of such cultural processes. Self-organization (e.g. de Boer, 2001, ch. 3) causes languages to remain uniform, even though variation is introduced continuously. Cultural evolution causes languages to remain usable and learnable, and historical contingency creates those aspects of language that can be explained as accidents of their history – properties and events that do not appear to have a clear functional explanation (an example of work that tries to identify whether patterns of language change are due to universal functional processes or historical contingencies is Dunn, Greenhill, Levinson, and Gray, 2011).

It is often a matter of debate whether biological adaptations are speech-specific or not (Evans & Levinson, 2009; Levinson & Evans, 2010). However, from a truly evolutionary point of view this is a meaningless question: as (biological) evolution needs to work with what is already there, everything involved in speech was used for something else before the advent of language and speech. The vocal tract was used for eating, breathing and non-linguistic vocal communication, while brain areas involved in speech and language may have been involved in fine motor control (Rizzolatti & Arbib, 1998) and communication (Taglialatela, Russell, Schaeffer, & Hopkins, 2008). Therefore, one could argue that no biological adaptation is truly specific to speech and language.[1] What *can* be investigated, however, is whether certain biological traits have undergone selection related to speech and language (although, because we are dealing with a very complex trait for which very little historical and comparative evidence is available, this is far from straightforward). Thus biological traits (both anatomical and cognitive) involved in speech and language can be divided into two classes: those that have changed because of selective pressure related to speech and language and those that have not – either because they have remained essentially the same since the latest common ancestor with other apes, or because they have changed for other reasons. Lively debate notwithstanding (Badin, Boë, Sawallis, & Schwartz, 2014; Boë et al., 2007; Boë, Heim, Honda, & Maeda, 2002; de Boer, 2010; Evans & Levinson, 2009; Hauser, Chomsky, & Fitch, 2002; Levinson & Evans, 2010; Pinker & Jackendoff, 2005), there is no consensus for any trait on whether it did or did not evolve because of speech and language.

A problem in determining which factors are responsible for speech and language is that there is no simple mapping between universal tendencies observed in languages and the mechanisms that are brought to bear in using and learning language (Kirby, Dowman, & Griffiths, 2007). In the study of universals of language, it has been assumed that properties that recur in languages tell us something about the underlying (cognitive or physical) mechanisms directly (Greenberg, 1963). Although it is true that properties of the language users must in some way determine the languages that they speak, it has been shown that the precise relation between learning biases and the languages that will emerge through cultural evolution is not necessarily straightforward. The nature of the relationship depends crucially on how learners generalize observed linguistic patterns in their own linguistic (re-)productions.

These conclusions have been derived from the mathematical analysis of Bayesian learning models. Griffiths and Kalish (2007) have investigated the case that learners reproduce the variation that they observe *exactly* according to the likelihoods that they derive from their observations and their prior biases. This behavior is called probability matching behavior. Languages that emerge in such a population will exactly reflect the learning biases of the learners. To give a simplistic example: if learners have a 80% preference for languages with final devoicing and a 20% preference for maintaining final voicing distinctions, then in 80% of the cases a language will emerge that has final devoicing. However, Kirby et al. (2007) have investigated the case in which learners tend to produce linguistic variants that they deem more likely more often than proportional to their likelihood (the most extreme case would be that they exclusively produce the most likely variant). In this case the language users *overgeneralize*. The languages that emerge in this case do not reflect the learners' biases directly. In fact, even very weak biases will result in the preferred languages dominating. For instance, the fact that one observes almost exclusively languages that use most of the available vowel space does not indicate that

[1] Unless one believes in macro-mutations that happen to give us complex behaviors, such as recursion (Chomsky, 2007). However, this would be unprecedented in biology.

Please cite this article as: de Boer, B. Biology, culture, evolution and the cognitive nature of sound systems. *Journal of Phonetics* (2015), http://dx.doi.org/10.1016/j.wocn.2015.07.001