Paleogenomics, hominin interbreeding and language evolution

Antonio Benítez-Burraco¹ & Lluís Barceló-Coblijn²

1) Departamento de Filología Española y sus didácticas, Universidad de Huelva, Campus de “El Carmen”, Avda. de las Fuerzas Armadas s/n, 21071 Huelva, Spain
   e-mail: antonio.benitez@dfesp.uhu.es

2) Universidad de Murcia, Departamento de Filosofía, Campus de Espinardo, E-30100 Murcia, Spain

Limited admixture between other hominins and anatomically-modern human (henceforth, AMH) populations has recently been confirmed. Non-African human genomes contain ≈ 1-4% of Neanderthal DNA (Green et al., 2010). Melanesian genomes additionally contain ≈ 4–6% of Denisovan DNA (Reich et al., 2010). Approximate-likelihood analyses of some Sub-Saharan African genomes are suggestive of the presence of ≈ 2% of archaic DNA, plausibly introgressed from an extinct hominin population that split from our ancestors about 700 kya (Hammer et al., 2011). It has been argued that AMH cognition could have been partially modelled by this kind of introgression events (Hawks et al., 2008). Linguistic abilities seem a natural target, in view of the fact that language has been traditionally regarded as a species-specific trait. A long-standing controversy divides the field concerning the possibility that language is a synapomorphic trait in late hominins, and particularly, that Neanderthals already had it (Mellars, 1996; d’Errico et al., 2003; Mithen, 2006). Genetic evidence has fuelled this debate. Different ‘language-related’ genes (i.e. genes that give rise to language disorders in our species when they are mutated) have been positively selected in our clade (e.g. Taipale et al., 2003, and Hannula-Jouppi et al., 2005 for some candidate genes for dyslexia). Crucially, Krause et al. (2007) found in Neanderthals the derived alleles of the two human substitutions in the FOXP2 protein (Enard et al., 2002), which seemingly predate to our common ancestor (who is also the ancestor of Denisovans), about ≈ 300-400 kya. FOXP2 is the ‘language gene’ par excellence (Vargha-Khadem et al., 2005; Fisher & Scharff, 2009; Benítez-Burraco, 2012). Not surprisingly, this find was regarded as the definitive piece of evidence for the Neanderthal language case (Trinkaus, 2007; d’Errico & Vanhaeren, 2009, p. 38; Frayer et al., 2010, p. 113). Can we be confident that the analysis of the DNA introgressed from archaic populations into AMHs actually sheds light on language evolution and the linguistic abilities of late hominins? We think that at this stage, caution is in order.

To begin with, all human languages appear to share some basic structural properties (Chomsky, 1980; Baker, 2001; Boeckx, 2009; among many others). Moreover, language acquisition is a process universal to the species (Lust, 2006; Slobin, 2006). Taken together, this suggests that all human beings are endowed with the same capacity for language. Given the human biological predisposition to acquire a language, the ‘linguistic genotype’ must be similar in all human beings as well. Consequently, since each AHM population incorporated different species-specific DNA portions, we should expect that this ‘linguistic genotype’ is not part of the introgressed DNAs. Moreover, this would imply that these species were endowed with it, and, ultimately, had modern, human-like language. However,
this is controversial, even for Neanderthals (see above). The opposite possibility is also appealing: extinct pre-modern hominins would have gained (or improved) their linguistic abilities when they interbred with AMHs. Although current analyses have only detected a gene flow from the former into the latter, but not vice versa, a reciprocal gene flow is certainly expected from the colonizing population to the resident population (see Green et al., 2010 for a discussion). Nonetheless, that possibility is also problematic. Firstly, evidence of modern language in other hominins is, once again, controversial. Secondly, the introgression events do not coincide with significant cultural changes. Allegedly it is modern language that fuels constant and feedback, AMH-specific cultural changes, because it allows to virtually explore new options and also to transmit the results in an efficient, quick, and instantaneous fashion (Dennett, 1995, 1996). As for Neanderthals, the admixture plausibly predates the emergence of modern-like cultures among them. Châtelperronian and related technocomplexes only emerged ≈ 40 kya (d’Errico et al., 1998; Langley et al., 2008). However, according to Green et al. (2010) the admixture took place ≈ 50-100 kya. In fact, they found that the exemplar from Mezmaiskaya, who lived ≈ 60-70 kya (Golovanova et al., 1999), is genetically similar to the specimens from Vindija and El Sidrón (we are not considering here the possibility that these ‘modern’ assemblages are not genuine innovations [Coolidge & Wynn, 2004; Mellars, 2005], or cannot be attributable to Neanderthals [Bar-Yosef & Bordes, 2010; Higham et al., 2010]). Thirdly, the introgressed DNA should have contained most (if not all) derived variants of the ‘language genes’ that were fixed after the split of our lineage from the line that gave rise to Neanderthals and Denisovans. The problem is that these genes are scattered throughout the whole genome (Smith, 2007; Benítez-Burraco, 2012). More importantly, for some of them we have direct evidence (see Green et al., 2010 supplementary material) that both species exhibited the ancestral alleles. For instance, Neanderthals were endowed with the ancestral alleles in some positions of ASPM and MCPH1, two genes related to brain size evolution (Zhang 2003; Evans et al., 2004). In the same vein, the Denisovan CNTNAP2 shows a fixed single nucleotide change compared to that of AMHs (Meyer et al., 2012). CNTNAP2 is one of FOXP2 targets (Vernes et al., 2008). It has been linked as well to specific language impairment (SLI) (Vernes et al., 2008), autism (Alarcón et al., 2008), and diverse clinical conditions in which language is disordered (Petrin et al., 2010; Sehested et al., 2010). Finally, the derived variants of all regulatory mechanisms of gene expression (and in fact, of the whole interactomes) should have been introgressed as well. The state (derived or ancestral) in other hominins of the regulatory networks of genes that are relevant for language is currently unknown. However, we know, for instance, that some of the physiological targets of FOXP2 (whose mutation gives also rise to language disorders) have been positively selected in our clade (Spiteri et al., 2007; Vernes et al., 2008; see above). Moreover, some differences have been attested as well concerning cis-regulatory regions of gene expression. For example, Maricic et al., (2012) have recently found an AMH-specific substitution within a regulatory region of FOXP2 which was known to have been affected by a selective sweep; this substitution is likely to alter FOXP2 expression via the transcription factor POU3F2. Eventually, as the analysis of the Neanderthal genome suggests (see Green et al., 2010 supplementary material for discussion), we expect hundreds of amino-acid sequence changes to be fixed in the AMH lineage after the divergence from Neanderthals and Denisovans, a greater number of potential regulatory substitutions, and also some fixed changes in human accelerated regions. It could even be the case that a differential activity of transposable elements or of viruses has modelled their respective genomes in dissimilar ways (Agoni et al., 2012). We cannot rule out the possibility that these specifically-human innovations have also played an important role in the development of neural devices involved in language processing.

Further, even if we could eventually prove that pre-modern hominins had the derived version
of the ‘linguistic genotype’ (partly due to these introgression events), a *Homo sapiens*-like linguistic phenotype is not granted. There is not a direct link between the genotype and the phenotype. Development is always synergistically regulated by multiple factors, which are all equally necessary (Oyama, 2000; Robert, 2008). Most of them remain substantially unknown in extinct hominins. Different genotypes can give rise to the same phenotype, but the other way around also holds (Pigliucci et al., 1996; West-Eberhard, 2003; Balaban, 2006). Neanderthals, Denisovans, and HAMs evolved in different environments (Finlayson, 2005; Carrión et al., 2011). Hence, the norms of reaction of their ‘linguistic genotypes’ could have been different as well. Notice also that genes are more important during the initial steps of development, when brains achieve a substantial degree of internal organization in advance of experience (Bouwman et al., 2004). However, their subsequent developmental trajectories (and plausibly the cognitive abilities they finally support) are modelled by other factors. Eventually, fully functional computational devices are only achieved when these pre-wired neural components are remodelled by the feedback effect exerted by the neural activity (inherent to language processing) (Balaban, 2006; Ramus, 2006). Importantly, Gunz et al. (2010, 2012) have found that the ontogenetic trajectories of the endocranium in AMH and Neanderthals differ most after birth, when these important changes in the wiring of the brain are taking place under the influence of environment (see also Petanjek et al., 2011). On another front, the real locus of selection is the whole set of phenotypes which make up entire organisms, not genomes (Sholtis & Weiss, 2005). Consequently, evolutionary novelties can arise in neutral conditions (i.e without genetic modifications) because of the dynamics and generative properties of developmental systems (Müller & Newman 2005; West-Eberhard, 2005). Modern language could be such an innovation.

In summary, introgressed DNA from archaic humans into AMHs (or vice versa) probably did not prompt the emergence of modern language. Other hominins could have had a ‘linguistic system’ and a similar ‘speech’. However, the available data suggests that the ‘languages’ they plausibly spoke would have lacked some defining properties of human languages, particularly, complex syntax, which is strongly based in recursive embedded structures, but exhibits as well other distinctive features, such as cross-serial dependencies (see Mellars, 1996 and Mithen, 2006 for similar conclusions regarding Neanderthal ‘language’). Other evolutionary mechanisms must account for its presence in our species only.

Acknowledgments

Antonio Benítez-Burraco’s research was funded by the Spanish Ministry of Economy and Competitiveness and FEDER through the grant ‘Biolinguistics: evolution, development, and fossils of language’ (FFI2010-14955). Lluís Barceló-Coblijn’s research was made possible through the grants ‘La naturalidad moral y estética humana. Caracterización sistemática de los rasgos derivados humanos de moral y estética’ (FFI2010-20759) and ‘Análisis corporizado de las nociones de computación, algoritmo e implementación y de la inducción de estructuras gramaticales’ (FFI2009-13416-C02-0), both funded by the Spanish Ministry of Economy and Competitiveness.

References


