

Evolutionary viewpoints on quantal vowels: A review of arguments for and against the existence of quantal vowels in *H. neanderthalensis*

Satu Hopponen
University of Eastern Finland

Abstract

During the past two decades, there has been considerable discussion on the existence of quantal vowels (as described by quantal theory) in the speech of an extinct hominin species, *H. neanderthalensis*. To put it simply, the argument is focussed on whether or not this species produced the kind of speech characteristic of *H. sapiens* today. In this article, the discussion is reviewed in detail, taking into account arguments for and against the notion. Finally, it is concluded that although the evidence is not conclusive enough to decide the matter once and for all, it seems likely that the *H. sapiens* -type vocal tract is a prerequisite for producing quantal vowels.

Keywords: speech, evolution, quantal theory, vocal tract

1 Introduction

This article is a review of the arguments for and against the existence of the capacity to produce quantal vowels in an extinct hominin species, *H. neanderthalensis* (hereafter HN). In order to keep this paper manageable, this article reviews only the material that is of direct relevance to the discussion at hand; exploring all the digressions would fill a small book. The first modelling attempt by P. Lieberman & Crelin was published in 1971. This view, though it had been challenged unsuccessfully by publications that did not include modelling attempts (for example, by Houghton 1993, which was rejected by P. Lieberman 1994), was challenged by another modelling attempt only three decades later (Boë et al. 2002). As P. Lieberman and Crelin had done before, Boë et al. (2002)

modelled the vocal tract of the same HN individual, La Ferrassie 1, but reached a very different conclusion. This is when the discussion became heated. The beginning of this century has brought with it a profusion of new papers on the topic. Despite the considerable length of the discussion, the core arguments of the proponents and opponents have remained essentially unchanged throughout, including the views P. Lieberman has defended the whole time.

The discussion is focussed around the three quantal vowels [i u ɑ], as originally described by Stevens (1972; 1989), which actually makes the quantal theory younger than the beginning of the discussion on HN's speech capabilities, when point vowels were modelled (P. Lieberman & Crelin 1971). However, quantal vowels seem to have been established as the main topic in the following discussion (see below). The questions of how, when and why the kind of speech that exists in modern humans today evolved are central to the discussion, since the answers offered by the proponents and opponents differ radically. Quantal vowels are important to *H. sapiens* speech because they make speech more robust (on error rates, see Peterson & Barney 1952; for a recent replication of Peterson & Barney's study, see Hillenbrand et al. 1995; for a description of quantal theory, see Stevens 1972; 1989; Stevens & Keyser 2010; for a recent review of the applicability of quantal theory to predicting whole vowel inventories, see, for instance, Diehl 2008).

The proponents and opponents can be divided into three rough categories: those who argue that HN could produce all three quantal vowels (Boë et al. 2002; Boë et al. 2007; Boë et al. 2013), those who argue against the existence of quantal speech in HN (P. Lieberman 2006a; 2007a; 2007b; 2012), and those who argue that HN could produce some quantal vowels (Barney et al. 2012). In addition, there exist other views, such as that the ear and vocal tract of modern humans is optimized for speech, as expressed by Martínez et al. (2013), who lean on the modelling studies done by Boë et al. (2002) when they claim that *H. heidelbergensis* was capable of producing quantal vowels (see below).

On the whole, the divide between proponents and opponents is quite clear-cut, which is rather unexpected since all of the arguments put forward are based on modelling the vocal tracts of extinct hominins. Where the arguments differ are the articulatory and acoustic results produced by these modelling attempts. The point of disagreement is described by de Boer (2009a: 257) as one concerning the limitations imposed by the vocal tract: the opponents see these limitations as the deciding factor, whereas the

proponents regard them as something that can be overridden by neural mechanisms. In previously published review literature, which is too vast to go through in this paper, reactions to the different arguments range from unanimous agreement with one party or another to cautious “wait and see” attitudes (for a short overview, see Hopponen 2014).

This paper is divided into three main sections. The first offers a short review of the discussion itself, focussing on the core arguments of the opponents and proponents. The second discusses the evolutionary explanations proposed for quantal vowels (and, to some extent, speech in general). The third and final section offers a few tentative conclusions. Although laid out in the papers reviewed in the first section, the argument over the possible vocal repertoire of *H. sapiens* infants is omitted. This is because its content is essentially identical to the HN discussion and because its inclusion would take up too much room (a review of it can be found in Hopponen 2014: 55–81).

For a detailed account on human evolution in general and HN in particular, the reader is referred to Cela-Conde & Ayala (2007). For a detailed account of the evolution of the human head and neck, the reader is referred to D. Lieberman (2011). For a slightly outdated but sharp in-depth critique of the proponents’ modelling attempts, see de Boer & Fitch (2010). For a paleontologically rather than linguistically oriented review of the modelling attempts, see Martínez et al. (2009). For an anthropological review on hominin vocal tracts, see Clegg (2012), and for a review focusing on the archeological side of the evolution of language and speech, see Lavento (2012). On the issue of whether language and/or speech are exclusive to humans or whether aspects of them can be found in other species, see, among others, Fitch (2010) for a review, and, in particular, Pepperberg (2002) for a detailed account of language and speech in African grey parrots.

2 The discussion so far

The different views are presented so that proponents (Boë et al. 2002; Boë et al. 2007; Boë et al. 2013) are followed by the opponents (P. Lieberman & Crelin 1971; P. Lieberman 2006a; 2007a; 2007b; 2012). They are then followed by the one publication whose results support the existence of a few quantal vowels (Barney et al. 2012), and Martínez et al. (2013), who reconstructed a *H. heidelbergensis* vocal tract but did not model any

vowels. Criticisms made by other researchers are scattered where appropriate.

Both proponents (Boë et al. 2002; Boë et al. 2007; Boë et al. 2013) and opponents (P. Lieberman & Crelin 1971; P. Lieberman 2006a; 2007a; 2007b; 2012) used the La Chapelle-aux-Saints skull (presumably, the individual number 1), but they resorted to different reconstructions of it. There is some minor disagreement over the significance of the different reconstructions; P. Lieberman (2007a: 556) denies that it has any significance, on the basis that the shape of the cranial base is not indicative of the shape of the vocal tract (the source he refers to is Fitch & Giedd 1999). Moreover, Boë et al. do not give any data or reasons in any of their publications regarding why the new reconstruction would be superior to the older one (on the ontogeny of *H. sapiens* cranial base, see Fitch & Giedd 1999 and D. Lieberman & McCarthy 1999; on the evolution of the cranial base, see, for instance, Bastir & Rosas 2016). The vertebrae of La Ferrassie were used to determine the length of the cervical and thoracic spine (P. Lieberman 2006a; 2007a: 466). Boë et al. (2007) specify that it was La Ferrassie 1's spine and the Kebara hyoid that were used. This illustrates the fact that the starting points for the different modelling attempts were rather similar, even though their results proved to be vastly different.

The existence of quantal vowels in HN has been explored by defining the maximal vowel space (e.g. Boë et al. 2002; Boë et al. 2007). This was done by mapping out three formants (F1, F2 and F3) (Boë et al. 2002), which are vitally important to quantal vowels (see, for instance, Stevens 1972; 1989). It is the view of Boë et al., expressed neatly in the title of their 2002 article, that “[t]he potential Neandertal vowel space was as large as that of modern humans”. Their view, which has remained unchanged throughout the discussion, is that HN could have used neural mechanisms to compensate for its long oral cavity (they assume this was neurally possible), and this compensation could have allowed enough manoeuvrability for quantal vowels to be produced (Boë et al. 2002). In a newer publication (Boë et al. 2007), they hold the view that neural mechanisms could compensate for a very ape-like vocal tract with a long oral cavity and a short pharynx. Strangely enough, Boë et al. (2013: 385), who do not model consonants in any of their studies, later attack their opponents for not taking consonants into account. In fact, consonants were modelled by P. Lieberman & Crelin (1971) in their original study, which concluded that HN may have been unable to produce velar plosives but might have been able to produce dental and labial plosives (P. Lieberman

& Crelin 1971: 216–217; on modelling consonants, see also Carré & Chennoukh 1995). The latest publication of the proponents is Badin et al. (2014), whose main conclusion is that so long as lips are included in the model, the height of the larynx is more irrelevant than de Boer (2010a; 2010b) claims (see below).

The assumption made by Boë et al. (2002), according to which neural mechanisms can override anatomical limitations, could be called *the compensation hypothesis*. This view has been sharply criticized by de Boer & Fitch (2010: 40–43), who reviewed the modelling attempts of Boë et al. as a case study. De Boer and Fitch (2010) write that since the model used by Boë et al. has *H. sapiens*-like vocal tract manoeuvrability built into it, of course the results then match the built-in capacities (see P. Lieberman 2012: 610 for similar criticism). De Boer & Fitch (2010: 43) conclude that this is a case of logical circularity. In addition, it has to be noted that the compensation hypothesis is close to being a null hypothesis since it appears to be unchallengeable; at present (and perhaps in the future as well), it is impossible to determine if HN had the neural capacity for fine motor control of the vocal tract.

There is one recent study that purports to support Boë et al.'s (2002) position. Martínez et al. (2013) claim that the proportions of the *H. heidelbergensis* vocal tract that they reconstructed are similar to those reconstructed by Boë et al. (2002). This conclusion Martínez et al. (2013: 96) reach just by reconstructing the vocal tract of Individual XXI of Sima de los Huesos and concluding that its proportions closely match the HN vocal tract that Boë et al. (2002) reconstructed and modelled. Martínez et al. do not model any sounds with their *H. heidelbergensis* vocal tract, and yet they claim that their results support Boë et al.'s (2002) results. Despite the fact that the ear modelling done by Martínez et al. (2013: 98–99) supports the notion that *H. heidelbergensis*' hearing was more similar to *H. sapiens*' rather than a chimpanzee's, merely the twin facts that the reconstructed bandwidth does not reach as high as *H. sapiens*' and that there is a considerable time differential between *H. heidelbergensis* and HN (on the possible family trees of these species, see, for instance, Mounier & Mirazón Lahr 2016) render Martínez et al.'s (2013) conclusion suspect. Further, as is the case with HN, we have no means by which to determine whether or not *H. heidelbergensis* possessed the neural compensation mechanisms proposed for HN by Boë et al. (2002).

The proponents' view is opposed by P. Lieberman and others (P. Lieberman & Crelin 1971; P. Lieberman 2006a; 2007a; 2007b; 2012), who

take the view that a vocal tract with 1:1 proportions is the *sine qua non* of quantal vowels. This is shown by the first publication (P. Lieberman & Crelin 1971: 209), in which the vocal tract of HN was modelled. They went so far as to skew the starting points, such as the position of the larynx, in favour of HN by giving it more *H. sapiens* -like characteristics. Despite this, the modelling showed that it is impossible to produce the full repertoire of sounds found in modern languages (P. Lieberman & Crelin 1971: 211–215). In much later studies, Carré & Chennoukh (1995) and Carré (2004) modelled the sounds produced by a *H. sapiens* vocal tract. In the latest study, Carré (2004: 238–239) came to the conclusion that the shape of the vocal tract that may be the limiting factor.

Further, it is impossible for HN to have a vocal tract with *H. sapiens* -like proportions (male or female) since attaining these proportions would place the larynx behind the sternum, which in turn would make it unique among hominins as well as unable to swallow (P. Lieberman 2006a: 301; 2007a: 557–558; P. Lieberman & McCarthy 2015). When discussing the standard shape of the mammalian tongue, P. Lieberman (2012: 613) writes that a flat tongue located in the oral cavity cannot descend far enough down the pharynx in order to produce [ɑ]. Since the only kind of vocal tract that we know to be capable of producing all three quantal vowels is the *H. sapiens* -type vocal tract, P. Lieberman's position is rather strong. In a much earlier study that focused on oral cavities only, Duchin (1990: 694) found that the oral cavities of HN, *H. sapiens* and *H. erectus* are similar – and thus differ from *Pan* – but it has to be noted that the study did not address the problem posed by the pharyngeal cavity. Thus, despite the fact that there are similarities in the shape of the oral cavity, the question of how an [ɑ] could be produced with a high larynx remains unsolved. At this point, it seems reasonable to suggest that the burden of proof rests on those who would claim that a high larynx is not an impediment to producing an [ɑ], since the only vocal tract that we know is certainly capable of it has a low larynx. Therefore, in light of the evidence put forward by P. Lieberman, Duchin and others, the claim made by Boë et al. rests largely on the compensation hypothesis, and therefore they have yet to produce overwhelming evidence to support their views.

There are others who share P. Lieberman's reservations, of course. De Boer & Fitch (2010) and de Boer (2010b; 2010a) have challenged Boë et al.'s (see above) conclusions. De Boer (2010a) modelled vocal tracts of different depths to determine the larynx height that bestows the greatest articulatory abilities, and found that the vocal tract of a female *H. sapiens*

to be optimal. The said modelling attempt has been accused by Badin et al. (2014) of oversimplifying things because the model does not include lips (which, in fact, it does include; see Section 5 of de Boer's paper for details). Badin et al.'s criticism seems to miss the mark slightly as it does not demonstrate that the stumbling block of neural mechanisms is removed from the path to quantal vowels. First, Badin et al. (2014) only modelled one lip aperture, and they did not take into account the enhancing gestures (described in Stevens & Keyser 2010), thus leaving it unsaid whether or not their model is able to distinguish between [i] and [y] or [u] and [ɯ], respectively. Second, they do not model the articulatory manoeuvres that are characteristic of quantal vowels (for a detailed account of these movements, see Stevens 1972; 1989; Stevens & Keyser 2010). Strangely, after making no mention of these rather important details, Badin et al. (2014: 167) explicitly claim that adding lips to the model validates what the proponents have been saying all along (they cite Boë et al. 2002; Boë et al. 2007; Boë et al. 2013). While Badin et al.'s (2014) criticisms are rather interesting and important, some clarification on the details is required.

The partially positive view is represented by Barney et al. (2012), who made their own (3-D) reconstruction of the vocal tract of HN. The only remarkable difference is that they used the spine of a modern human (a large man's) instead of a HN spine (Barney et al. 2012: 92) and that they used both female and male modern humans when they modelled the vocal tract they used for reference (Barney et al. 2012: 91). When they had built the HN vocal tract, they modeled the vowel space with F1 and F2 (Barney et al. 2012: 92), leaving out the rather important F3 (see, for example, Stevens 1972; 1989). The result of Barney et al.'s modelling attempt is that their HN could not produce an [a] – which, it has to be pointed out, is not a quantal vowel – with an F1 that is comparable to data collected from modern humans (Barney et al. 2012: 98). Further, they opine that their results are not comparable with previous studies (Barney et al. 2012: 97). They are very cautious in their conclusions and write that their attempt should be regarded as hypotheses testing (Barney et al. 2012: 100). This sort of caution seems reasonable enough, especially when compared to the – at times – aggressive manner in which the proponents and opponents have put forth their views. Further, it is interesting and refreshing that Barney et al. (2012) did not approach the matter from the point of view of any of the existing modelling attempts but attempted to redo the modelling from scratch.

Barney et al.'s attempt at modelling is problematic in a few ways, however. The fact that the cartilages of the larynx have to fit into the neck (P. Lieberman 2006a: 301; 2007a: 557; in addition, see P. Lieberman & McCarthy 2007 and P. Lieberman & McCarthy 2015 for details on relative vocal tract lengths) is not mentioned by Barney et al. (2012). Since they used the spine of a large *H. sapiens* male (large in terms of weight or height? They mention no numbers), it is likely that the cartilages fit into the neck despite the longer oral cavity. Because in *H. sapiens* the cervical spine is rather long (P. Lieberman 2013: 144–145; P. Lieberman & McCarthy 2015), it is unclear whether or not the cartilages would fit into a HN neck. It might have been better that Barney et al. (2012) had used a HN spine, such as the La Ferrassie spine (according to D. Lieberman 2011: 538, La Ferrassie 1 consists of both a skull and a spine; further, P. Lieberman 2013: 145 writes that there exist well-established methods for measuring fossil spines; see P. Lieberman & McCarthy 2015: Table 2 on ventral heights of cervical vertebrae; see also D. Lieberman 2011: 588–589 on palate lengths and vocal tract proportions in both HN and *H. heidelbergensis*; see also Bastir & Rosas 2016 on the changes in the cranial base and face during the evolution of *Homo*).

Another point that Barney et al. (2012) do not address is that in order to enable swallowing, the tongue has to fill the oral cavity (D. Lieberman 2011: 297). However, there are no established methods to estimate the size of an extinct hominin's tongue because it is not clear if the relationship between body size and tongue size is isometric or not (D. Lieberman 2011: 335–336), and therefore the omission in Barney et al. (2012) is understandable. Nevertheless, together these points leave one wondering how realistic the vocal tract of HN modelled by Barney et al. (2012) actually is. Further, it seems that this model (Barney et al. 2012) might suffer from similar logical circularity as Boë et al.'s (described above), since the vocal tract that was used as reference was constructed using modern humans – although in this case, the possible circularity follows from the materials used rather than the assumed flexibility of the vocal tract.

Further, it might be beneficial to keep in mind the criticisms presented in other publications. For instance, de Boer (2009a) is of the opinion that, at least in modern humans, a female vocal tract is capable of producing a greater range of formants than the male one. He also writes that since other evolutionary pressures, such as size exaggeration, may have influenced the male vocal tract, it might be wiser to use the female vocal tract as a

template when modelling extinct hominins (de Boer 2009a: 264–265) (on the ontogeny of female and male vocal tract, see, for instance, Fitch & Giedd 1999). On discussing the historical aspect of this, Clegg (2012: 65) writes that the reason that a male vocal tract was originally used has to do with the weight of history and accepted practice. In light of this, one cannot help but wonder what kind of results the modelling attempts of P. Lieberman and Boë et al. would produce if they used the remains of female hominins instead of males (and if Barney et al. had used only females in the construction of the reference vocal tract). In addition, it remains to be seen how much anatomy can ultimately tell of the function(s) of different structures, including the different parts of the vocal tract (see, for instance, D. Lieberman 2008).

There exist two recent reviews of the evolution of speech that, from a linguistic point of view, miss the mark rather widely. Lavento (2012) is an archeological review that practically rules out linguistics as a central participant in the discussion and, judging from its references section, does not seem to be aware of the last four decades of discussion between the proponents and opponents. An equally recent anthropological review of the speech capabilities of HN is Clegg (2012: 66, 73), which does not review the linguistic side of the matter beyond circa 1992 but which nevertheless claims that HN did not “have a restricted range of speech sounds” because the overall length of the HN vocal tract was within the range observed in the modern human females. She does not specify which sounds – consonants, vowels or just quantal vowels – she means. Clegg (2012: 66) even admits that she has not modelled the sounds that the HN vocal tract might have produced. Her argument, though cogent and likely supported by the anthropological evidence that she cites, misses the mark by a wide margin. The debate on the linguistics side of the divide between disciplines has not been about vocal tract length itself but about vocal tract shape and manoeuvrability.

It is very problematic that similar modelling methods produce contradictory results (de Boer 2009a: 256), especially since all of the studies used nearly the same set of HN fossils. This probably is the result of faulty or inadequate modelling methods, but it also could be the result of not having established methods for estimating the proportions and shape of vocal tracts on the basis of bony landmarks alone. Or, indeed, this state of affairs could well be the result of not having satisfactory methods with which to determine the range of movements that those modelled vocal tracts are capable of performing. Therefore, it seems reasonable to assume

that individual muscles – and their insertions and shapes – will need to be modelled in order to determine the range of movements that any given hominin would have been able to execute (on the differences in tongue musculature in chimpanzees (*Pan*) and humans, see Duchin 1990; among others, Takemoto 2001 has already attempted to describe musculature of the human tongue in the context of modelling it). It may be that much more data from extant apes is required, since at least *Australopithecus afarensis* had the hyoid bulla (Alemseged et al. 2006: 298), if the vocal tracts of even earlier hominins are to be modelled in the future. Additional modelling attempts, focussed on more hominin species than is currently the case, might shed more light on the emergence of quantal vowels and, indeed, speech itself.

3 Evolutionary explanations

When discussing the evolution of language, it is important to steer clear of unfalsifiable (null) hypotheses (see, for instance, D. Lieberman 2011: 587–588). Happily, if one leaves aside the possible logical circularities, this kind of unproductive argumentation has largely been avoided in the arguments for and against the existence of quantal vowels in the speech of HN.

The obvious difference between the explanations offered by the proponents and the opponents in the argument over speech capacities of HN is that the opponent, P. Lieberman (2007b: 52) can offer an evolutionary reason for the life-threatening shape of the human vocal tract: in his view, the risk is offset by the benefits of modern human speech. This view has been countered by Clegg (2012: 68–69), who is of the opinion that the risk of choking on food is minimal. It is unfortunate that the issue has not yet been looked at from a cross-species perspective, since there might be something to be gained by looking at the rate of death by choking, for example, in either of the chimpanzee species or in other apes. It is obvious, however, that such data would not be easy to collect, and therefore data from other mammalian species with high larynges might serve.

Further, P. Lieberman (2013: 144–145) proposes that since it is difficult to fit a vocal tract with 1:1 proportions into a short neck, the longer neck of modern humans contributes to our ability to produce quantal vowels. De Boer (2009a: 256) makes a similar point when he writes that, in the case of the vocal tract of modern humans, the benefits counterbalance the drawbacks (for a similar view, see de Boer 2010a).

In addition, de Boer (2012) opines that vocal communication was the reason why hominins lost the hyoid bulla and the air sacs (on modelling the effects of air sacs in conjunction with a *H. sapiens* vocal tract, see de Boer 2009b). This is supported by a histological study, which analyzed the Kebara 2 hyoid (from a male HN), that comes to the conclusion that some measurements, but not all, of the hyoid fell within the range documented for modern humans (D’Anastasio et al. 2013). Cautiously, they conclude that their results mean that the hyoid participated in similar activities as the hyoid of a *H. sapiens*, but they also point out that this does not mean that HN was able to speak (D’Anastasio et al. 2013: 6). At least once, it has been suggested that the shape of the hyoid dates back to the origins of the genus *Homo* (Clegg 2012: 71–72), which would, in Clegg’s opinion, imply that the human-like vocal tract is equally old. On a more cautious note, she does write (Clegg 2012: 71) that more data (i.e. more hyoids) is needed. Thus, there are no cast-iron certainties to be found on the hyoid front either, yet.

At a glance, all of the studies and hypotheses described in section 2 seem to be within the realm of possibility, and it is easy to assume that the development and/or emergence of speech may have played a significant role in the evolution of the vocal tract. It has to be noted, however, that other (possibly complementary) explanations, some from outside the domain of linguistics, have also been proposed; D. Lieberman et al. (2001: 124–125) suggest that it is possible that swallowing sets more restrictions on the shape of the vocal tract than quantal speech does (for additional explanations and caveats, see, for instance, D. Lieberman 2011: 415). In a similar vein, Nishimura (2005: 202) writes that the most basic functions of the vocal tract, namely eating and breathing, can set limits to the optimization of the vocal tract for speech. An identical point is made by Carré (2004: 239). Recently, it has been suggested by Coquerelle et al. (2013: 4–7) that the protrusion of the chin seen in modern humans could be the result of how, among other things, the larynx descends and the facial block rotates during early infancy. Further, they suggest that the need to swallow and breathe safely could act as a driving force for change (Coquerelle et al. 2013: 2). It has also been suggested by Clegg (2012: 66–69, 74) that the particulars of human speech sounds are the product of accident and that there is little reason to assume that the shape of the human vocal tract is dangerous enough to exact a price, contra the opponents’ views. Because these questions fall outside linguistics proper,

they are best left for other disciplines to answer, but it seems reasonable to assume that linguistic input may be required at some point.

The lack of evolutionary explanations in the publications of Boë et al. (Boë et al. 2002; Boë et al. 2007; Boë et al. 2013) is obvious. If they continue to insist that a vocal tract with 1:1 proportions is not necessary for quantal speech and that articulatory compensation (together with neural mechanisms) could have overcome anatomical obstacles for quantal vowels, it would be interesting if they could offer an alternative evolutionary explanation for the curiously shaped vocal tract of *H. sapiens*. In fact, they insist that the modifications to the vocal tract need not have preceded speech (Boë et al. 2002: 481). This is slightly confusing, but in keeping with their view. Needless to say, the complementary explanations offered by D. Lieberman and Nishimura (see above) do not mesh with the views of Boë et al. as well as they do with P. Lieberman's.

However, any account of the evolution of human language and speech will need to take into account the fact that other species seem to share parts of these abilities (for the abilities of, for instance, African grey parrots, see Pepperberg 2002; for a review of the language and speech capacities of other nonhuman species, see Fitch 2010). At least one other primate species, the Diana monkey (*Cercopithecus diana*) has been the topic of a disagreement over attempts at vocal tract modelling; in chronological order, the publications that discuss it in detail are Riede et al. (2005), P. Lieberman (2006b), Riede et al. (2006), P. Lieberman (2007b) and P. Lieberman (2013). There is relatively little novelty in this argument, and participants, with the opponents' side taken again by P. Lieberman and Riede enacting the part of the proponents, argue for the same things as in the HN debate. The only thing missing from it are the quantal vowels. However, since *Cercopithecus diana* is only distantly related to modern humans and our ancestors, this discussion seems more like a sidetrack than anything else, especially since the quantal vowels are absent. Its importance to the discussion at hand is that it shows that there is something to be gained from studying extant species, even if the only result that comes out of it is yet another disagreement.

4 Conclusions

Given the content of the discussion, it might be better to dub the discussion *Arguments for and against quantal vowels in extinct hominins*, since, at present, neither the proponents nor the opponents support the notion that

extinct hominins were incapable of producing some kind of speech (which would, presumably, include a vowel or vowels). Aside from P. Lieberman & Crelin (1971), there have been no attempts to define what sort of consonants were possible; de Boer (2009a: 260) writes that the methods used in exploring vowels are insufficient to determine possible consonants. Further, evolutionary explanations for what caused the shape of the vocal tract to change so dramatically in the hominin lineage remain elusive.

It is troubling that previous modelling attempts (described in section 2) have reached such different conclusions. Based on their results, it seems reasonable to assume that quantal vowels were out of the reach of HN, but the question cannot, at least not without further modelling attempts and a synthesis of additional supporting evidence (outlined in Section 3), be considered as settled. As mentioned above, it is problematic that modelling attempts with very similar starting points – and even the same fossils – have reached different conclusions. From this, it would be very tempting to draw the conclusion that all modelling attempts have somehow failed, but surely such a sweeping conclusion would not only be wrong but also do injustice to the researchers involved. Regardless, the problem of contradictory results persists, and until further modelling attempts are made, there are no grounds for declaring the issue settled once and for all.

There is nothing to suggest that quantal vowels are a necessary feature of speech (P. Lieberman 2007b: 41) or, indeed, (proto)language. It is conceivable, although strictly hypothetical, that the number of vowels could have been very low indeed, perhaps so low as to comprise only a front–back or low–high dichotomy (Jussi Niemi, p.c. 2013-03-27; cf. Stevens' description of these dichotomies 1989: 15). Further, there is no reason to assume that quantal vowels had to develop simultaneously (which would yield some kind of 'partial quantality'). Therefore, results similar to those of Barney et al. (2012) might be expected. Aside from the results of the modelling studies described above, it is far from certain that a specific set of (quantal or other) vowels existed at some point in the distant past. However, P. Lieberman (2007b: 41) considers quantal vowels to be an innovation. The fact that HN and *H. sapiens* coexisted for a long time (Cela-Conde & Ayala 2007) does complicate matters somewhat. It is possible that a) one species had quantal vowels, b) both had them, or c) neither had them. The first two alternatives are supported by the modelling studies described above, while the third, although logically possible, is supported by no empirical evidence at all. The matter is complicated

further if ‘partial quantality’ is thrown into the mix. In the end, it is all about timing.

It is noteworthy that the current trend is not to specify a point in time when a particular feature of language or speech emerged (cf., for instance, Bickerton 1990). The more branches the human family tree sprouts, the wiser this seems; at present, there is limited certainty as to which species leads to which (on family trees and how to attempt to draw them, see, among others, Mounier et al. 2009; Cela-Conde & Ayala 2007). Thus, it seems prudent to avoid drawing conclusions on the capabilities of any particular species on the basis of (preliminary) studies done on other species (see above).

The fact that quantal vowels are so ubiquitous in the languages spoken by modern humans does require an explanation, however. The shape of the vocal tract is a very likely explanation, but the structure of the ear has most likely played a role as well (Martínez et al. 2013). It would be entirely speculative to ponder the exact point in time when hominins were neurally capable of producing and perceiving quantal vowels – in other words, when they could reap the benefits offered by these vowels.

It is clear that the disagreement over the existence of quantal vowels in the speech of extinct hominins cannot be resolved satisfactorily by resorting to linguistic methods only. That any discipline could, by itself and without needing input from others, solve the puzzle seems unlikely, and therefore the problem remains a multidisciplinary one. That the attempts to resolve the issue have drawn on many disciplines is heartening and has generated much discussion. Any future modelling attempts would probably greatly benefit from the expertise that, for instance, paleontology, archeology, computer science and evolutionary biology can offer. Hopefully, in the future, some light can be shed on the emergence and development of quantal vowels. One such possible avenue of research could be offered by optimality theory (for a review, see Parker & Maynard Smith 1990), but attempting such an explanation is far beyond the scope of this review.

Finally, it can be concluded that although both the proponents and the opponents of the core argument have put forward much evidence, the balance of evidence suggests that it might be wise to err on the side of caution and conclude that there is no conclusive evidence in favour of the existence of quantal vowels in HN. One can only hope that future attempts at solving the issue will be more successful in mapping the articulatory

capacities of extinct hominins, shedding light on evolution of the human vocal tract and finding out how and why quantal vowels emerged.

Acknowledgements

I would like to thank the anonymous reviewers for their insightful comments, which improved this manuscript considerably. I would also like to thank my PhD supervisors, Stefan Werner and Alexandre Nikolaev, for their support and encouragement.

References

- Alemseged, Zeresenay & Spoor, Fred & Kimbel, William H. & Bobe, René & Geraads, Denis & Reed, Denné & Wynn, Jonathan G. 2006. A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature* 443(7109). 296–301.
- Badin, Pierre & Boë, Louis-Jean & Sawallis, Thomas R. & Schwartz, Jean-Luc. 2014. Keep the lips to free the larynx: Comments on de Boer's articulatory model (2010). *Journal of Phonetics* 46. 161–167.
- Barney, Anna & Martelli, Sandra & Serrurier, Antoine & Steele, James. 2012. Articulatory capacity of Neanderthals, a very recent and human-like fossil hominin. *Philosophical Transactions of the Royal Society B* 367(1585). 88–102.
- Bastir, Markus & Rosas, Antonio. 2016. Cranial base topology and basic trends in the facial evolution of Homo. *Journal of Human Evolution* 91. 26–35.
- Bickerton, Derek. 1990. *Language and species*. Chicago: University of Chicago Press.
- Boë, Louis-Jean & Heim, Jean-Louis & Honda, Kiyoshi & Maeda, Shinji. 2002. The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics* 30(3). 465–484.
- Boë, Louis-Jean & Heim, Jean-Louis & Honda, Kiyoshi & Maeda, Shinji & Badin, Pierre & Abry, Christian. 2007. The vocal tract of newborn humans and Neanderthals: Acoustic capabilities and consequences for the debate on the origin of language. A reply to Lieberman (2007a). *Journal of Phonetics* 35(4). 564–581.
- Boë, Louis-Jean & Ménard, Lucie & Captier, Guillaume & Davis, Barbara & McNeilage, Peter & Sawallis, Thomas R. & Schwartz, Jean-Luc. 2013. (Letter to the editor) Anatomy and control of the developing human vocal tract: A response to Lieberman. *Journal of Phonetics* 41(5). 379–392.
- Carré, René. 2004. From an acoustic tube to speech production. *Speech Communication* 42(2). 227–240. doi:10.1016/j.specom.2003.12.001.
- Carré, René & Chennoukh, Samir. 1995. Vowel-consonant-vowel modeling by superposition of consonant closure on vowel-to-vowel gestures. *Journal of Phonetics* 23(1–2). 231–241.
- Cela-Conde, Camilo & Ayala, Francisco J. 2007. *Human evolution: Trails from the past*. Oxford: Oxford University Press.

- Clegg, Margaret. 2012. The evolution of the human vocal tract: Specialized for speech? In Bannan, Nicholas (ed.), *Music, Language, and Human Evolution*, 58–80. New York: Oxford University Press.
- Coquerelle, Michael & Prados-Frutos, Juan Carlos & Rojo, Rosa & Mitteroecker, Philipp & Bastir, Markus. 2013. Short faces, big tongues: Developmental origin of the human chin. *PLoS ONE* 8(11). 1–9.
- D’Anastasio, Ruggero & Wroe, Stephen & Tuniz, Claudio & Mancini, Lucia & Cesana, Deneb T. & Dreossi, Diego & Ravichandiran, Mayoorendra & Attard, Marie & Parr, William C. H. & Agur, Anne & Capasso, Luigi. 2013. Micro-biomechanics of the Kebara 2 hyoid and its implications for speech in Neanderthals. *Plos ONE* 8(12). 1–7. doi:10.1371/journal.pone.0082261.
- de Boer, Bart. 2009a. Why women speak better than men (and its significance for evolution). In Botha, Rudolf & Knight, Chris (eds.), *The prehistory of language*, 255–265. (Studies in the Evolution of Language 11). Oxford: Oxford University Press.
- 2009b. Acoustic analysis of primate air sacs and their effect on vocalization. *Journal of the Acoustical Society of America* 126(6). 3329–3343.
- 2010a. Modelling vocal anatomy’s significant effect on speech. *Journal of Evolutionary Psychology* 8(4). 351–366.
- 2010b. (Letter to the editor) Investigating the acoustic effect of the descended larynx with articulatory models. *Journal of Phonetics* 38(4). 679–686.
- 2012. Loss of air sacs improved hominin speech abilities. *Journal of Human Evolution* 62(1). 1–6.
- de Boer, Bart & Fitch, W. Tecumseh. 2010. Computer models of vocal tract evolution: An overview and critique. *Adaptive Behavior* 18(1). 36–47.
- Diehl, Randy L. 2008. Acoustic and auditory phonetics: The adaptive design of speech sound systems. *Philosophical Transactions of the Royal Society B* 363(1493). 965–978. doi:10.1098/rstb.2007.2153.
- Duchin, Linda E. 1990. The evolution of articulate speech: Comparative anatomy of the oral cavity in Pan and Homo. *Journal of Human Evolution* 19(6). 687–697. doi:10.1016/0047-2484(90)90003-T.
- Fitch, W. Tecumseh. 2010. *The evolution of language*. Cambridge: Cambridge University Press.
- Fitch, W. Tecumseh & Giedd, Jay. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America* 106(No. 3, Pt. 1). 1511–1522.
- Hillenbrand, James & Getty, Laura A. & Clark, Michael J. & Wheeler, Kimberlee. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5, Pt. 1). 3099–3111.
- Hopponen, Satu. 2014. Homininien ääntöväylärekonstruktioiden tarkastelua kvantaaliteorian valossa [An overview of hominin vocal tract reconstructions in the light of quantal theory]. Joensuu: University of Eastern Finland. (MA Thesis.)
- Houghton, P. 1993. Neandertal supralaryngeal vocal tract. *American Journal of Physical Anthropology* 90(2). 139–146.

- Lavento, Mika. 2012. Puheen evoluutio – arkeologin näkökulma [The evolution of speech – An archeologist’s perspective]. In Aaltonen, Olli & Lehtinen, Mona & Aulanko, Reijo (eds.), *Näkökohtia puheen alkuperästä* [Perspectives on the origin of speech]. Tampere: Mediapinta.
- Lieberman, Daniel E. 2008. Speculations about the selective basis for modern human craniofacial form. *Evolutionary Anthropology* 17(1). 55–68.
- 2011. *The evolution of the human head*. Belknap Press, Cambridge: Belknap Press of Harvard University Press.
- Lieberman, Daniel E. & McCarthy, Robert C. 1999. The ontogeny of cranial base angulation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions. *Journal of Human Evolution* 36(5). 487–517.
- Lieberman, Daniel E. & McCarthy, Robert C. & Hiiemae, Karen M. & Palmer, Jeffrey B. 2001. Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of Oral Biology* 46(2). 117–128.
- Lieberman, Philip. 1994. Functional tongues and Neanderthal vocal tract reconstruction: A reply to Houghton (1993). *American Journal of Physical Anthropology* 95(4). 443–452. doi:10.1002/ajpa.1330950408.
- 2006a. *Toward an evolutionary biology of language*. Cambridge: Belknap Press of Harvard University Press.
- 2006b. Limits on tongue deformation – Diana monkey formants and the impossible vocal tract shapes proposed by Riede et al. (2005). *Journal of Human Evolution* 50(2). 219–221.
- 2007a. Current views on Neanderthal speech capabilities: A reply to Boe et al. (2002). *Journal of Phonetics* 35(4). 552–563.
- 2007b. The evolution of human speech: Its anatomical and neural bases. *Current Anthropology* 48(1). 39–66.
- 2012. Letter to the Editor: Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics* 40(4). 608–622.
- 2013. *The unpredictable species: What makes humans unique*. Princeton: Princeton University Press.
- Lieberman, Philip & Crelin, Edmund S. 1971. On the speech of Neanderthal man. *Linguistic Inquiry* 2(2). 203–222.
- Lieberman, Philip & McCarthy, Robert C. 2007. Tracking the evolution of language and speech: Comparing vocal tracts to identify speech capabilities. *Expedition* 49(2). 15–20.
- Lieberman, Philip & McCarthy, Robert C. 2015. The evolution of speech and language. In Henke, Winfried & Tattersall, Ian (eds.), *Handbook of paleoanthropology*, 873–920. 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Martínez, Ignacio & Quam, Rolf & Arsuaga, Juan Luis & Lorenzo, Carlos & Gracia, Ana & Carretero, José-Miguel & Rosa, Manuel & Jarabo, Pilar. 2009. Approche paléontologique de l’évolution du langage: Un état des lieux. *L’anthropologie* 113(1). 255–264.
- Martínez, Ignacio & Rosa, M. & Quam, Rolf & Lorenzo, Carlos & Bonmatí, Alejandro & Gómez-Olivencia, Asier & Gracia, A. & Arsuaga, Juan Luis. 2013.

- Communicative capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Quaternary International* 295. 94–101.
- Mounier, Aurélien & Marchal, François & Condemi, Silvana. 2009. Is *Homo heidelbergensis* a distinct species? New insight on the Mauer mandible. *Journal of Human Evolution* 56(3). 219–246.
- Mounier, Aurélien & Mirazón Lahr, Marta. 2016. Virtual ancestor reconstruction: Revealing the ancestor of modern humans and Neandertals. *Journal of Human Evolution* 91. 57–72.
- Niemi, Jussi. Personal communication. 2013-03-27.
- Nishimura, Takeshi. 2005. Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *American Journal of Physical Anthropology* 126(2). 193–204.
- Parker, G. A. & Maynard Smith, John. 1990. Optimality theory in evolutionary biology. *Nature* 348(6296). 27–33.
- Pepperberg, Irene M. 2002. *The Alex studies: Cognitive and communicative abilities of Grey Parrots*. Cambridge: Harvard University Press.
- Peterson, Gordon & Barney, Harold L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24(2). 175–184.
- Riede, Tobias & Bronson, Ellen & Hatzikirou, Haralambos & Zuberbühler, Klaus. 2005. Vocal production mechanisms in a non-human primate: Morphological data and a model. *Journal of Human Evolution* 48(1). 85–96.
- Riede, Tobias & Bronson, Ellen & Zuberbühler, Klaus. 2006. Multiple discontinuities in nonhuman vocal tracts -- A response to Lieberman (2006). *Journal of Human Evolution* 50(2). 222–225.
- Stevens, Kenneth N. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In David, Edward & Denes, Peter B. (eds.), *Human communication: A unified view*, 51–66. (Inter-University Electronics Series Vol. 15). New York: McGraw-Hill.
- 1989. On the quantal nature of speech. *Journal of Phonetics* 17(1–2). 3–45.
- Stevens, Kenneth N. & Keyser, Samuel Jay. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38(1). 10–19.
- Takemoto, Hironori. 2001. Morphological analyses of the human tongue musculature for three-dimensional modeling. *Journal of Speech, Language, and Hearing Research* 44(1). 95–107.

Contact information:

Satu Hopponen
School of Humanities, Philosophical Faculty
University of Eastern Finland
PO Box 111
80101 Joensuu
Finland
e-mail: shoppone(at)student(dot)uef(dot)fi