

## 6SSEL045 – Language Origins

## Lecture 5

## Sound and Gesture

Human language uses several physical systems which are peculiarly human: the vocal tract; the auditory system; facial expression; complex gesture; and, of course, the unique human brain. These systems did not all evolve as a single package. This lecture looks at the likely when, who and how of the evolution of the species-specific mechanics behind language.

**THE VOCAL TRACT**

The vocal tract consists of:

- **The lips, cheeks and frontal mouth parts.** These developed early in the evolution of vocal communication – chimpanzees use lip smacks, kissing noises, plosives and nasals in their social sound-making. These sounds seem to be, at least partially, under volitional control. For instance, orang utans use a volitional kiss-squeak call to warn a predator that it has been seen, and to warn any other orangs in the area that a predator is around (Lameira & Call, 2018).
- **Bipedality.** It may seem odd to include this as part of the vocal tract, but walking on two legs gives greater control over breathing. Quadrupeds need to synchronise their breathing with the movement of their front legs; bipeds do not (Provine & Yong, 1991). This means that our vocalisation is unaffected by our ambulatory activity – we can walk and talk at the same time. However, this still doesn't necessarily mean we are good at performing both actions simultaneously (Hyman et al, 2010).
- **The chest cavity, lungs, and air flow.** These are key components of the chimpanzee pant-hoot call, but the chimpanzee has less control over their air flow than humans, partly because they have retained the capacity to breathe and swallow at the same time (Belin, 2006). This is one of the reasons why the attempts to teach Gua and Viki to talk didn't succeed: humans have greater control over air flow, and therefore over our range of vocalisations.
- **The tongue & vocal chords.** Control of these seems to have developed relatively late in our evolution. For chimpanzees, the tongue and vocal chords just add volume and tone; the main articulators for sound signalling are the chest cavity and lungs (over which they also have less control than humans). Australopithecines are unlikely to have had good control of the tongue and vocal chords, merely because the motor systems in their brains were not as large as in the Homo clade. Early Homo probably had greater control, but by how much? We cannot know for sure because tongues & vocal chords are soft tissue and do not fossilise. There is some indication of the range of movements possible from the muscle anchor points in the jaws of fossils, but the vocal actions actually used remain open to interpretation. We have only one good clue to the amount of vocal control in early Homo: in both Homo sapiens and Neanderthals, the hyoid bone is reduced, which indicates greater muscle control in the vocal chords (Nishimura et al, 2006). This may mean that the common ancestor of Neanderthals and humans (possibly Homo erectus), could have spoken in a languagelike way; but this is still a disputed view.

Why the human vocal tract developed in the way it did is also an issue. One proposal is that it developed for singing (Mithen, 2005). Singing is a signal used by many species, including apes: it is used to indicate personal fitness (the main reason for birdsong), as a way of indicating where you are (the main reason for solo calls), as an indicator of group size (the main reason for chorusing), and to build and maintain social relationships (the probable reason for antiphonal singing – Jordan et al, 2004). The costly signalling approach says that singing to indicate fitness drives an evolutionary race toward song complexity – the more complex the song, the fitter the individual, and the more breeding success they have (Locke, 2017). This is

supported by a recent study into the highly variable and multimodal displays of birds of paradise (Ligon et al, 2018): male birds with a superior display (which often includes a vocal element) mate with more females.

Some researchers claim that the vocal tracts of chimpanzees and humans are too different to support a simple route from common-ancestor vocalisation to human language, so they propose a gestural stage of prehuman communication (e.g. Arbib et al, 2008). There is, however, sufficient time in the evolution of Australopithecus and Homo (about 170,000 generations) for an incremental vocalisation route. In addition, both modern chimpanzee and modern human communication systems incorporate both vocalisation and gesture; it is therefore likely that vocalisation and gesture evolved together. This is explored in more detail below.

**THE AUDITORY SYSTEM**

The receiving end of signalling is often forgotten, but there are important differences between chimpanzee and human hearing systems (Quam et al, 2012; Belin, 2006). There is even some evidence of difference at the genetic level: "The gene with the most significant pattern of human-specific positive selection is alpha tectorin, whose protein product plays a vital role in the tectorial membrane of the inner ear." (Clarke et al, 2003, p1962). This gene seems to affect acuity of hearing, and possibly tonal differentiation. Genetic differences affecting hearing have also been found between H. sapiens and Neanderthals, although it is less clear what these differences signify (Gómez-Olivencia et al, 2015).

It seems likely, therefore, that the auditory changes leading to language occurred continuously over the whole evolutionary period from the common chimpanzee-human ancestor to H. sapiens. There was not a single mutation that made our hearing system speech-friendly, there was a cumulative, non-neutral effect driving us toward better speech perception.

Recognition of individual signallers is an important feature in the development of language: the meaning of a vocalisation is not just about the message, the sender and the receiver are significant contexts in the meaning of the message: for the receiver, the meaning of the signal involves both the message and the sender. This semantic multiplicity seems to be an ancient feature of communication in the primate clade: other primates do not just hear a noise and respond, there is evidence that they relate the signal to the signaller, and respond appropriately depending on context (Engh et al, 2006).

**FACIAL EXPRESSION**

Charles Darwin (1872) thought that there were six core human expressions which seem to have universal meaning: anger, fear, disgust, happiness, sadness, and surprise. He saw these expressions as non-volitional – we produce them because of our emotional state. However, there are other facial signals which seem to have universal form and meaning, like laughter, the yawn of boredom, the eyebrow flash of recognition, the furrowed forehead of concentration, and the single raised eyebrow of scepticism, and these can be volitional. There are also some facial expressions (often deceptive or group-based signals) which are culturally defined (e.g. the sideways glance can have different meanings, depending on culture). Different cultures also read different parts of the face to interpret expressive meaning (Gendron et al, 2014).

We now recognise four universal expressions: happiness; sadness; fear/surprise; and anger/disgust. The expressions for fear and surprise, and for anger and disgust, seem to be culturally interchangeable (Jack et al, 2012). It used to be believed that human facial expression was much more communicative than that of our nearest relatives. We now know that, in terms of musculature, we are very similar. Chimpanzee expressiveness may appear to us to be

limited, but that is largely because the differences in bone structure make the muscle effects different (Burrows et al 2006) – and it is also because, while they use similar expressions to those we use, they do not always mean the same thing. Most notably, the chimpanzee exposure-of-teeth expression may look like a smile; but it is, like the dog exposure-of-teeth display, a sign of anger or fear.

### COMPLEX GESTURE

Chimpanzees cannot produce complex gesture, because it relies in large part on the ability to walk bipedally, freeing the hands (Schmidt, 2003); but they do have a range of kinaesthetic signals. Hobaiter & Byrne (2014) have identified at least 66. Many seem to be genetic conventions (the same across all groups); but a small number, like the Mahale grooming handclasp (McGrew et al, 2001), are group-specific and therefore cultural conventions (and probably symbolic). This contrasts with humans, where most gestural signals are cultural and therefore symbolic, and liable to very different interpretations in different cultures. This is not because humans have fewer non-volitional signals, we have largely the same range as chimpanzees; it is because we have many more volitional signals. Because our gesture system is mostly cultural and volitional, humans have a gestural channel of communication which can be as rich as speech; this is why deaf sign languages are full languages.

### LANGUAGE IN THE BRAIN

Apart from size, and therefore capacity, there are few visible differences between human and chimpanzee brains. The chimpanzee has a brain size of about 400cc (the bonobo brain size is similar to australopithecines, at 350cc), but almost all the bumps (gyri) and dips (sulci) in the cortex are similarly located in all three species (Herculano-Houzel, 2009). The traditional language areas (Broca's and Wernicke's areas, the angular and supra-marginal gyri, and the auditory and motor cortices) are present in all three brains, and they seem to work in the same way – which indicates that, in humans, their primary role may not actually be linguistic.

There are some differences between the brains that are not immediately obvious: the neuronal architecture of human brains seems to be more complex, with a greater number of interneuronal connections than in the other two brains (Mattson, 2014). The more tightly-packed architecture of human brains may also give them a processing-speed advantage, although this has not been definitively shown to be the case.

There, are though, two areas where human brains are more developed than Pan brains.

- The orbitofrontal cortex seems to be involved in defining the social context of the self (Azzi et al, 2012; Watanabe et al, 2018), and it is underdeveloped in the chimpanzee brain when compared to the modern human brain. This seems to indicate that chimpanzees do not have as detailed a sense of self as humans do. Humans are particularly adept at treating themselves as if they were other people – that is, being disinterested in their own self. This allows modern humans not just to empathise but to sympathise with others, and also to self-sacrifice (Edwardes, 2014).
- The fronto-polar cortex seems to be involved in maintaining partial solutions in memory while other partial solutions are generated, so that all the partial solutions can be integrated into a general solution (Green et al, 2015). It therefore seems to be linked to the capacity to produce iterative and hierarchical thought, which are both important considerations in the generation of language constructs. The FPC is more developed and more complex in humans than in chimpanzees. However, the exact role of the FPC is disputed.

Despite the apparent geographical specialisations in the human cortex, Exner's area shows that there is a problem with over-modularising the human brain. Sigmund Exner thought that the area above Broca's area was implicated in writing and reading problems.

We now know that there has been insufficient time since the appearance of scripting systems for the brain to evolve innate processes and areas for reading and writing skills (Roux et al, 2010), and for those genetic effects to become widespread in the species. Instead, writing seems to plug into general pattern-recognition systems which have a wide variability in the human population – hence the high incidence of agraphia and dyslexia.

### WHAT HAPPENED WHEN?

While we cannot know for sure how the mechanics of language came to be in place, we can make an educated guess about the timescale.

- A large portion of the mechanics of language were already in place in the chimpanzee-human common ancestor.
- Bipedality was the first important difference, and it seems to have evolved about 6 million years ago (Richmond & Jungers, 2008). Early Australopithecines were already functionally bipedal, but only the Homo clade seems to have had the necessary airflow control for effective use of speech (Provine, 2004).
- Human facial expression and complex gesture seem to have evolved about 2 million years ago, with *Homo ergaster* and *Homo erectus* (Barnard et al, 2007).
- The human labio-laryngeal system (tongue and larynx) is old in evolutionary terms; but the cognitive capacity to use them for speech seems to have evolved about 1 million years ago, with *Homo heidelbergensis* (Fitch, 2010).
- Two significant increases in brain size seem to be implicated in language evolution: the jump from 700cc to 900cc with *Homo ergaster/erectus*; and the jump from 900cc to 1250cc with *Homo heidelbergensis* (Antón et al, 2014). The first is likely related to complex gesture, the second is likely related to the reorganisation of the labio-laryngeal (vocal) system. However, it is unlikely that they were the causes of the increases in brain size, just useful side-effects.

### WAS HUMAN LANGUAGE ORIGINALLY GESTURAL?

Complex gesture is tied to a whole series of capacities: tool-making, tool use and music-making, to name a few (D'Errico et al, 2003); but where does this place complex gesture in the evolution of humans? Australopithecines were bipedal and used tools (Ward et al, 2011), and the later Australopithecines also made tools (Sayers et al, 2014); so was complex gesture and dexterity an early route to complex communication?

Many people believe so, and they suggest that modern language emerged from a gestural form of communication (e.g. Arbib, 2008; Corballis, 2002, 2010; Frey, 2008; Gillespie-Lynch et al, 2014; Kendon, 2017; Pollick & de Waal, 2007; Sterelny, 2012). It is certainly true that today, even excluding deaf sign languages, speech has a major body language component: gestural communication is as important to us as it is to chimpanzees.

There is a problem with this gestural route to language, however: if chimpanzees use both gestural and vocal signalling, and they use them contextually (Crockford & Boesch, 2003), then why, during the development of the human species, should there be a conversion to a mainly gestural form of communication, and then another conversion back to a mixed method? It seems like an overly complex route with no good reason to take it: if you already have the advantages of both types of signalling, why abandon them? To do so would imply that the signalling environment of early humans was sufficiently different from those of both modern chimpanzees and modern humans; this may be true, but there is no evidence to suggest this is the case.

We used to believe that chimpanzees have no, or very limited, volitional vocalisation; but we now know this is not so: they can use (and do use) volitional vocal communication (Watson et al, 2015).