Commentary/Ackermann et al.: Brain mechanisms of acoustic communication in humans and nonhuman primates

Abstract: Ackermann et al. mention the “acquisition of species-atypical sounds” in apes without any discussion. In our commentary, we demonstrate that these atypical sounds in chimpanzees not only include laryngeal sounds, but also have a major significance regarding the origins of language, if we consider looking at their context of use, their social properties, their relations with gestures, their lateralization, and their neurofunctional correlates as well.

Whether apes are able to voluntarily and intentionally control their vocal production remains a topic of intense debate (e.g., Hopkins et al. 2011). In a brief paragraph in their target article (sect. 2.1.4.), Ackermann et al. mention the “observational acquisition of species-atypical sounds” in apes and acknowledge that chimpanzees are able to produce voluntary sounds using the modulation of the air through the lips (“blowing raspberries” or “kiss”). However, the authors also claimed that apes are not able to “engage laryngeal sound-production mechanisms” that can be “decoupled volitionally from species-typical audiovisual displays.” In fact, this latter claim is not accurate.

Hopkins et al. (2007) have indeed described the use of two atypical novel “learned” sounds produced by several chimpanzees among the captive groups from the Yerkes Primate Research Center: Some chimpanzees are not only able to produce non-voiced “raspberries” or “kiss” sounds (involving only the lips with the air of the mouth) but also “extended grunts,” which clearly engage the vocal tract and laryngeal sound-production mechanisms. Hopkins and colleagues showed that the production of these atypical sounds and vocalizations is often produced with pointing gestures and is used exclusively in the presence of both a human and an out-of-reach food in order to beg for food, while typical species-specific “food calls” were more frequent in the presence of food alone (Hopkins et al. 2007). Such atypical productions were interpreted as signals used intentionally to capture the attention of the human. Indeed, great apes have been shown to use those acoustic signals—vocal and lip sounds, cage banging or clapping gestures—especially when the recipient is not attentive, whereas visual pointing gestures are preferentially used when the recipient is attentive (e.g., Leavens et al. 2004; 2010; see also in orangutans: Cartmill & Byrne 2007; for a review of the literature, see Hopkins et al. 2011). In other words, the multimodal flexibility of communicative signaling (sounds, vocalizations, and gestures) is a manifestation of the ability of the great apes to adjust the modality of the signal to the attentional state of the recipient, and such an intentional property might be thus a special feature of social cognition that is needed in language processing.

In addition, given the inter-individual variability among chimpanzees concerning the ability to produce or not those novel sounds, it has been interpreted that, as for human speech but in contrast to species-typical vocalizations, those atypical vocal and lip sounds might be socially learned. In fact, it has been reported that chimpanzees raised by biological mothers who were able to produce those sounds, were more likely to also be able to do so than chimpanzees raised by humans in a nursery (Taglialatela et al. 2012). Moreover, among the chimpanzees that were not able to produce these atypical vocalizations, a recent study not
only showed that (i) it was possible to explicitly train them to do so using operant conditioning, but also (ii) that those subjects would further use these novel vocalizations in a communicative context for getting the attention of a human (Russell et al. 2013).

Finally, the investigation of lateralization of those atypical sounds and its functional cerebral correlates show some continuity with the language system. Indeed, most of the language functions involve a left-hemispheric dominance (Knecht et al. 2000). Interestingly, it turns out that these chimpanzee auditory signals, when produced simultaneously with food-begging pointing gestures, induce a stronger right-hand preference than when the gesture is produced alone (Hopkins & Cantero 2003), indicating that the left hemisphere may be more activated when producing both gestures and these atypical vocal and lip sounds simultaneously. Moreover, measures of orofacial asymmetries for vocal production in chimpanzees have showed that species-typical vocalizations – such as food barks or pant-hoot – elicited a left-sided orofacial asymmetry (i.e., right-hemispheric dominance), whereas atypical attention-getting sounds elicited an asymmetry toward the right side of the mouth, indicating that, as for right-handedness for communicative clapping gestures (Meguerditchian et al. 2012), a left-hemispheric dominance might be involved for producing those acoustical signals (Losin et al. 2008). More impressively, brain imaging studies (PET [positive emission tomography]) conducted in three captive individuals have found that communicative signaling for begging food from a human by using either gestures, atypical attention-getting sounds, or both of these modalities simultaneously, activated a homologous region of Broca’s area (IFG) predominantly in the left hemisphere (Tagliatela et al. 2008), a pattern of activation which is enhanced in subjects who used both gestural and vocal signals simultaneously (Tagliatela et al. 2011).

These collective findings support the idea that the atypical orofacial and vocal sounds in chimpanzees are a good illustration of the potential existence of a multimodal intentional system that integrates gestures, orofacial, and atypical vocal sounds into the same lateralized system. This multimodal communicative system not only shares some features of social cognition and social learning with human language, but also seems to be ultimately related to brain specialization for language (Meguerditchian et al. 2011). This theory is consistent with the evidence that in humans, a single integrated communication system in the left cerebral hemisphere might be in charge of both vocal and gestural linguistic communication (e.g., Gentilucci & Dalla Volta 2008). For all of these reasons, and their implications for the precursors of human language and its brain specialization, we believe that Ackermann et al. should better consider these voluntary laryngeal sound-production mechanisms in chimpanzees and the related multimodal communicative system, in their theoretical model.

ACKNOWLEDGMENT
This research was supported by a grant from Agence National de la Recherche ANR-12-PDOC-0014-01 (LangPrimate).