Semantic abilities evolved as well - Electronic commentary on M. Arbib: 'From monkey-like action recognition to human language'

Jean-Louis Dessalles, Laleh Ghadakpour

HAL Id: hal-00616377
https://hal.science/hal-00616377

Semantic abilities evolved as well

Jean-Louis Dessalles
ParisTech ENST (CNRS UMR 5141)
dessalles@enst.fr

Laleh Ghadakpour
LTCI, CNRS (UMR 5141)
laleh@ghadakpour.net

Abstract: The evolutionary story proposed in the target paper makes no difference between semantic representations underlying language and more general cognitive representations, at work in perception and action, which humans share with apes and probably other mammals. Though semantic representations supporting language are grounded in perception, some of them, specifically predicative structures, should rather be considered a distinctive feature of human communication system. Any evolutionary scenario about language should explain how human minds evolved to form the kind of thoughts that are communicated through language.

In his attempt to explain the presence of language in our lineage, M. Arbib makes generous assumptions about pre-hominid cognition and yet the transition to language remains enigmatic. One of the most challenging issues about language is why human beings, like bees, but unlike other primates, spend much time and energy in referential communication (DESSALLES 2000). The most blatantly unsupported presupposition in the target paper is the author's belief in some general biological urge to transmit skills and useful information to unrelated conspecifics, which non-human primate species, due to some cognitive limitations, would have been unable to satisfy. The author's concern is thus limited to finding some cognitive bifurcations that merely enabled human communication.

Even on this ground, the scenario imagined by M. Arbib seems unnecessarily intricate. In a first step, our ancestors are supposed to master a pantomime-like analog code, relying on imitation capabilities. Then, in a second step, their followers developed the ability to associate conventional holistic signs to frequently encountered scenes. In a last non-biological step, some sapiens invented compositional languages. The most obvious problem with this scenario is that the middle, conventional step undermines any necessity for the initial, imitative stage. Why did not hominids associate conventional signs with frequent situations in the first place? Association is intrinsically reversible, and does not require any complex mirroring machinery to serve as a code. Moreover, conventional signs are claimed to fulfil a disambiguation purpose. Why didn't we evolve more powerful analog imitating abilities, instead of regressing to mere holistic signalling?

Our strongest scepticism concerns the fact that the cognitive bifurcations postulated by M. Arbib do not affect semantic abilities. Because of their various cognitive skills, apes and hominids are generously granted with the same semantic representations as modern sapiens: they are supposed to recognise, decompose and assemble cognitive schemas, the kind of schemas that they are able to imitate. Such abilities appear both too rich and too poor to account for the evolutionary emergence of language.

Monkey brains are claimed to host hierarchical predicative structures such as GRASP(LEO,RAISIN). The purpose of a language evolution scenario is thus merely to explain how "the ability to communicate in a way that reflects these structures" could develop. However, predicative structures are relevant to language, and not to perception

(DESSALLES & GHADAKPOUR 2003). Though our brain is able to simultaneously detect a variety of complex perceptive structures in some given visual scenery, predicative structures like GRASP(LEO,RAISIN) are *constructed* for communicative purposes from the perceived scene rather than being a mere translation of it. One obvious reason is that we can systematically express the negative version of a predicative structure, e.g. "Leo doesn't grasp the raisin", while there is no perceptive meaning corresponding to the negation of a visual scene (what would be the negative version of Picasso's *Guernica*?). If semantic representations could be translated indifferently into language or into actions, as the author assumes, we would expect something like negative actions, as there are negative sentences. But this does not exist.

Our claim is that the ability to form predicative representations is part of the language faculty, and that action-oriented perceptive representations, as rich as they may be, are not what language communicates in the first place. Arbib's proposal that some actions like grasping or running can be directly communicated through mimetic gestures would be attractive if the mechanism could be extrapolated to the whole range of communicable objects, relations and judgements. But there is no natural imitation of an object like raisin. Drawings on paper, on cave walls or in the air are too conventionalised to be a direct translation of semantic representations. Similarly, while language offers words for relations, imitation can only attempt to show instantiated situations. Contrary to spoken or signed phrases, mimetic performances for '*on* the ground', '*on* the table', '*on* my hand' bear no systematic resemblance. This is unexpected if semantic relations are supposed to shape our perceptions and actions. The same can be said of judgements like 'a huge raisin' or 'a huge tree', which we would be bound to imitate in a systematic manner if something like HUGE(X) was represented in our mind regardless of any communication.

Predicative structures lie neither in scenes nor in gestures. Though some authors see in perceptive mechanisms interesting prerequisites for predicate formation (HURFORD 2003), genuine predicates cannot be reduced to perceptive structures (DESSALLES & GHADAKPOUR 2003). *They are constructed to be communicated*. When a raisin is predicated as huge, it is not merely perceived as such. It is *contrasted* with the speaker's prototypic raisin, and the adjective 'huge' happens to be the closest word associated with that contrast (GHADAKPOUR 2003). During this contrasting procedure, a large number of perceptive details of the scene are ignored. Predicates appear to be cognitive constructs, which exist for one main purpose: language. They are much poorer than perceptive representations, but they can bear propositional attitudes like assertion, disbelief, worry, etc. (DESSALLES 2000).

The purpose of human communication is not to create an accurate copy of a scene in the addressee's mind through some iconic or symbolic code. Language is not a poor substitute for video reporting. Most scenes are too rich and the expressive powers of language are too narrow and to slow. Moreover, the reason why addressees would be interested in comprehensive descriptions of scenes is still missing. Arbib is certainly right to try to ground the expressiveness of language in perception and in action. A sound scenario of the evolutionary emergence of language, though, would not strive to explain how hominids managed to make accurate descriptions, be it through pantomime or through words. Such a scenario should rather explain *why* and *how* our ancestors got minds able to form predicative structures, and to express them through compositional languages.

References

JEAN-LOUIS DESSALLES, 2000, *Aux origines du langage : une histoire naturelle de la parole*, Hermès.

JEAN-LOUIS DESSALLES & LALEH GHADAKPOUR, Object Recognition Is Not Predication (commentary on "The Neural Basis of Predicate-Argument Structure" by J. Hurford), *Behavioral and Brain Sciences*, 26(3): 290-291.

LALEH GHADAKPOUR, 2003, *Le système conceptuel, à l'interface entre le langage, le raisonnement et l'espace qualitatif : vers un modèle de représentations éphémères* (Thesis Dissertation), École Polytechnique.

JAMES R HURFORD, 2003, "The Neural Basis of Predicate-Argument Structure", *Behavioral and Brain Sciences*, 26(3): 261-283.