cortex 48 (2012) 923–935

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Viewpoint

Can syntax appear in a mirror (system)?☆

Marco Tettamanti [a,b,*] and Andrea Moro [a,c]

[a] Division Neuroscience, San Raffaele Scientific Institute, Milano, Italy
[b] Department of Nuclear Medicine, San Raffaele Scientific Institute, Milano, Italy
[c] Institute for Advanced Study IUSS, Pavia, Italy

ARTICLE INFO

Article history:
Received 31 March 2010
Reviewed 7 June 2010
Revised 23 August 2010
Accepted 20 May 2011
Published online 1 June 2011

Keywords:
Language
Mirror neuron system
Phonetics
Semantics
Syntax
Hierarchical

ABSTRACT

Converging evidence indicates that the processing of some aspects related to the phonetic and the semantic components of language is tightly associated with both the perceptual and the motor neural systems. It has been suggested that mirror neurons contribute to language understanding by virtue of a neurophysiological response matching perceptual linguistic information onto corresponding motor plans. This proposal has sometimes been extended to advocate that the language competence as a whole, including syntax, may be ascribed to this kind of perceptuo-motor mappings. This position paper examines what kinds of empirical and theoretical challenges such general mirror neuron language accounts need to face in order to proof their validity – challenges that we think have not been adequately addressed yet. We highlight that the most important limitation is constituted by the fact that some core defining properties of human language, at the phonetic, semantic, and especially at the syntactic level, are not transparent to the bodily senses and thus they cannot be the direct source of mirror neuron perceptuo-motor matching.

1. On the relation between language and the motor system

As the body of research in support of a tight link between language processing and the cortical motor system rapidly grows (for recent reviews see Willems and Hagoort, 2007; Fischer and Zwaan, 2008), so does the literature against the view that the neurobiological mechanisms ascribed to the Mirror Neuron Systems (MNS; Rizzolatti et al., 2008) provide exhaustive explanatory power for understanding the neural bases of language (Toni et al., 2008; Lotto et al., 2009; Hickok, 2009). MNS-based language theories posit that, in the course of hominid evolution, a specific class of perceptuo-motor neurons – so called "mirror neurons" – has incorporated the capacity to respond to communicative speech gestures. By analogy with the action execution-observation matching mechanism mediating action understanding, it has been speculated that the MNS may contribute to the understanding of communicative meaning (Rizzolatti and Craighero, 2004). In this theoretical position paper, we will intentionally avoid any commitment concerning the hypothesis that language evolved from a more rudimentary combination of sensorimotor and cognitive functions [but see Tettamanti et al. (2009) for a detailed discussion of this issue]. In particular, we will not argue against, nor in favor of the conjecture that linguistic communication evolved from manual gestures, as suggested

☆ Author contributions: although the authors certainly share all opinion expressed in this paper, Marco Tettamanti takes full responsibility for the ideas expressed in Section 2, and Andrea Moro for those expressed in Section 3.

* *Corresponding author*. Scientific Institute San Raffaele, Via Olgettina 58, I-20132 Milano, Italy.
E-mail address: tettamanti.marco@hsr.it (M. Tettamanti).

0010-9452/$ – see front matter
doi:10.1016/j.cortex.2011.05.020

by Corballis (2010) among others, or other type of motor control, as suggested by Toni et al. (2008) proposing that the evolution of language may be consequent to the emergence in humans of voluntary control over the vocal apparatus. Rather, we will argue that a single neurophysiological mechanism, such as the essential property of the MNS to map together perceptual and motor information, is inherently unable to explain all the manifold components of the human language system (for similar views, see Arbib, 2010; Corballis, 2010).

Research in this field has been driven by two main independent classes of hypotheses: one relies on the idea that the phonetic segments of speech are recognized by the listener by mapping the intended phonetic forms onto the articulatory motor programs involved in the production of the corresponding speech sounds; the other posits that the understanding of the semantics of actions described linguistically involves, at least in part, the mental simulation of the motor programs associated to the very same actions. As we will describe in some more details in the next section, both the phonetic and the semantic MNS hypotheses emphasize the direct and mandatory role in language understanding of the mapping of speech sounds and meaning respectively to the corresponding motor programs. More specifically, language understanding[1] is supposed to be, at least in part, mediated by perceptuo-motor mappings that have evolved from the neurophysiological audio-motor and visuo-motor properties of mirror neurons demonstrated in non-human primates (Gallese et al., 1996; Kohler et al., 2002). The hypothesis of sensorimotor-dependent interpretive processes for language phonetics and semantics is also shared with several other, independent theoretical accounts, namely those based on neuropsychological evidence of semantic category-specific deficits (Warrington and McCarthy, 1987), on hebbian-rule-like neurophysiological mechanisms (Pulvermüller, 2008), or on a combination of the two (Martin, 2007; Patterson et al., 2007). Most of these other theoretical accounts rely on associative mechanisms linking semantic or phonetic neural representations to congruent, anatomo-functionally distinct sensorimotor representations (which may also include the MNS), on the basis of experience. However, in our view, a fundamental aspect distinguishes MNS theories from associative accounts. In the latter, the understanding of linguistic sounds or meaning is thought to be accomplished by the fast reverberation of neural activity within experience-dependent, distributed neural networks involving sensory and motor representations [with or without the intervention of higher-order integration centers, see Patterson et al. (2007)]. In other words, language understanding is mediated by functional/effective connectivity among distributed brain regions (again, possibly including the MNS). In turn, MNS theories pose the emphasis on non-associative mechanisms, where the understanding of linguistic sounds and meaning is (at least partially) accomplished through the intrinsic neurophysiological properties of a single class of cells, i.e., mirror neurons: in mirror neurons sensory information triggers an internal motor resonance evoking action meaning on the basis of the available motor vocabulary (di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 2001; Rizzolatti and Craighero, 2004). Note that, given these different properties of the two accounts, the mapping of perceptual linguistic information onto the sensorimotor system may under some circumstances play just an accessory role in an associative model (e.g., because the perceptual information is not fully manifest to the bodily senses or because motor action is not particularly relevant for a given linguistic item), whereas it becomes necessary for language understanding if one assumes a non-associative MNS model (see also Lotto et al., 2009). In this latter view, if, for any reasons, the perceptuo-motor mapping cannot occur in mirror neurons, then the resonance-based extraction of relevant information needed for understanding does not take place: under such circumstances, understanding fails altogether. Also note that in order for the MNS to capture linguistic meaning, there must be a one to one equivalence between the intended linguistic structure and the sensory information perceived by the bodily senses, something which may not always be the case. In fact, as we will argue, the hypothesis that the phonetic and the semantic structure of language are mainly processed by the MN perceptuo-motor system is in principle sound, but with some notable caveats. In addition, although the phonetic and the semantic MNS hypotheses share the same idea that the motor system plays a major role, clearly, the two hypotheses refer to quite different linguistic and neuropsychological mechanisms. With this respect, current proposals do not sufficiently motivate how a single class of perceptuo-motor neurons can accommodate both mechanisms.

The picture described so far, however, is further complicated by the radical idea that has been sometimes advanced – although to the best of our knowledge never thoroughly tested – that the MNS accommodates the interpretive mechanisms for other linguistic components, such as hierarchical syntactic structures (Pulvermüller and Fadiga, 2010). With respect to language syntax, there are major theoretical reasons to argue against this view. The main argument that we will defend in our article is precisely that, since some of the core structural properties of syntax are not directly accessible to hearing and vision, as well as to any other bodily senses, an MNS-based account of the structural properties of syntactic competence, or more radically a unitary MNS theory of the relationship between language and the motor system, is not tenable. We can anticipate the detailed argument, by noticing that this conclusion is implied by the very nature of hierarchical syntactic relations. Indeed, according to most widely accepted generative grammar approaches [for a review, see Jackendoff (2003)], complex syntactic structures result from combinatorial operations assembling words into two-dimensional hierarchical relations that can be established at long distance following well-defined boundaries and constraints. These boundaries and constraints, as well as long-distance hierarchical relations become largely invisible to perceptual systems when the trees are compressed into linear, i.e., mono-dimensional, strings of words, such as in speech (Chomsky, 1995; Kayne, 1994; Moro, 2000). In fact, this operation of linearisation implies that syntactic processing must also rely on

[1] By "understanding" we very broadly mean the process of capturing the relevant structural information that allows us to correctly interpret communication at a given linguistic level, e.g., at the. phonetic level, the proper recognition of the word *be* as opposed to *bee*; at the semantic level, the proper representation of the intended meaning, e.g., *bee* versus *fly*, etc.