# Can syntax appear in a mirror (system)?

Marco Tettamanti[1,2] and Andrea Moro[1,3]

*[1]Division Neuroscience and [2]Department of Nuclear Medicine, San Raffaele Scientific Institute, Milano, Italy; [3]Institute for Advanced Study IUSS, Pavia, Italy.*

Corresponding author:

Marco Tettamanti, PhD, Scientific Institute San Raffaele, Via Olgettina 58, I-20132 Milano, Italy, Tel ++39-02-26434888, Fax ++39-02-26434892; Email: tettamanti.marco@hsr.it

Manuscript in press within the Cortex special issue on “Language and the motor system”

Total number of characters (including spaces): 76'980

Total number of words: 11'541

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 are 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 can not 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 Haagort, 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 favour of the conjecture that linguistic communication evolved from manual gestures, as suggested 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 (Pulvermueller, 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

[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* vs. *fly*, etc.

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 can not 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 (Pulvermueller 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 different computational capacities than just those based on perceptuo-motor mappings or associations. Nevertheless, even hidden syntactic properties may under some circumstances interact with the MNS: syntax interacts with both phonology and semantics during language comprehension and production and, hence, complex (i.e. not simply in terms of internal motor resonance as postulated by MNS theories) interactions between syntax and the MNS may be observed: as we will illustrate, this may in fact be the case for sentential negation.

## 2. On sound and the lexicon: what is likely to be mirrored in language.

In this section we will focus on what we think that a proper linguistic source of stimulus for the MNS can be, namely some aspects of speech sound perception and some aspects of lexical semantics pertaining to predicates expressing action. As we will point out, however, also in the phonetic and lexical-semantic domains, just as in the syntactic domain, there are core defining properties that are largely unattainable to perceptual systems.

### *2.1 Perception of speech sounds*

Based on psychoacustic evidence, the motor theory of speech perception (Liberman and Mattingly, 1985; Galantucci et al., 2006) proposed that the speech phonetic segments (importantly, not the associated articulatory gestures perceived by vision) are directly, i.e. without preliminary acoustic decoding, mapped onto the corresponding motor representations required for articulating the same speech sounds. This original proposal has been re-elaborated within an MNS theory approach (Rizzolatti and Craighero, 2004), predicting that perceptuo-motor mirror neurons in the premotor cortex, including Broca's area, should respond both when perceiving and when producing a speech sound. One of the key findings that led to this hypothesis was that in area F5 of the macaque ventral premotor cortex there is a class of so called audio-motor mirror neurons that appear to access the meaning of actions through the sounds that these actions produce. Accordingly, in addition to hand action execution, audio-motor mirror neurons discharge following the presentation of a hand action-related sound, but not following the presentation of a meaningless sound (Kohler et al., 2002).

The phonetic MNS hypothesis was tested in humans by measuring modulations of the motor evoked potential (MEP) of tongue muscles evoked by TMS of the corresponding somatotopic region of the motor cortex, when perceiving Italian words. These words could contain either linguo-palatal fricative consonants that involve strong articulatory movements of the tongue (e.g. *bi*[*rr*]*a*), or labio-dental fricative consonants that do not require such movements (e.g. *ba*[*ff*]*o*). As expected, stronger MEPs were found for "tongue words" compared to "non-tongue words". These facilitatory effects can be induced by TMS as early as 100 ms after the onset of a critical speech sound. The same effects, although reduced in amplitude, were found for tongue vs. non-tongue pseudowords. This suggests that the lexicality of the stimulus has no major influence on motor cortex activity, but leaves open the possibility that lexicality has some reinforcing effects (Fadiga et al., 2002). Kotz and colleagues (2010) elaborated on this notion, combining behavioral, TMS and fMRI to study phonological rhyme priming with either words or pseudowords. Based on their results, they proposed that, while the ventral premotor/motor cortex may underlie a low-level mapping of speech sounds onto motor representations that is independent of lexical meaning, Broca's area may underlie a higher-order process in speech perception that is only engaged in the presence of meaning. Thus, contrary to the original motor theory of speech perception, speech sounds *per se* are not invariantly mapped onto motor representations; rather, mapping takes place to a greater extent when the perceived speech sounds convey lexical meaning.

This body of evidence has more recently received support from a set of fMRI studies, where listening to meaningless syllables was shown to activate the same portions of the motor and premotor cortex that are also activated by producing the same syllables (Wilson et al., 2004). Furthermore, these overlapping activations reflect the somatotopic organization of the motor and premotor cortex (see Buccino et al., 2001; Postle et al., 2008) when the syllables investigated started with phonemes that involve movement of the lips (e.g. [*p*]) or the tongue (e.g. [*t*]) (Pulvermueller et al., 2006). The causal contribution of these somatotopic activations was demonstrated with TMS (D'Ausilio et al., 2009).

In sum, convergent evidence has demonstrated a link between language processing at the phonetic level and the motor system. However, the overall significance of these data does not seem to be completely clear. Lotto and colleagues (2009), for example, noticed that the motor theory of speech perception does not correlate with evidence coming from aphasic patients: lesions involving the motor speech areas typically do not compromise language comprehension, contrary to expectation. This may suggest that the kind of associations reviewed above are better interpretable in terms of modulatory sensorimotor associations within a wider distributed neural system for speech processing, rather than in terms of a focal, non-associative perceptuo-motor mapping involving just one class of neurons, i.e. mirror neurons. This alternative interpretation is also supported by the results of Kotz and colleagues (2010), given that some kind of integration process, at a higher level than simply sensorimotor, is proved to be necessary to determine a differential response of Broca's area to either word- or pseudoword-phonemes. A further hint in favor of distributed associations linking speech production to speech perception comes from proposals suggesting the need for neural activity associations between Broca's area, mirror neurons in the ventral premotor cortex coding for speech sound motor programs and the auditory cortex, to provide auditory feedback control of speech production (e.g Guenther et al., 2006; Tourville et al., 2008; Golfinopoulus et al., 2010).

Finally, the view that the intended phonetic structure of language and the produced/perceived acoustic speech stream are linked by a one to one relationship is oversimplified. This is indicated for instance by the phenomenon of coarticulation, i.e. the modulatory influence exerted on a given phonemic segment by the preceding and following segments (Toni et al., 2008; Lotto et al., 2009). In order to correctly distinguish and interpret the perceived speech sounds, any forms of direct perceptual mapping are not sufficient. The listener's brain must perform some kind of context-dependent processing that integrates the phonemic or intended phonetic segments across forward and backward-going time windows. In this sense, the view that phonetic information is directly accessible and fully transparent to the bodily senses is not correct, and the phonetic MNS hypothesis fails to provide a comprehensive theoretical account. This reinforces the perplexities on the pervasive role of the MSN in language processing addressed in this paper.

### *2.2 Comprehension of action-related lexical semantics*

There have been numerous independent scientific reasons leading to the claim that the organization of conceptual knowledge in semantic memory closely reflects the association with the dominant sensorimotor channels of experience for a given concept (Warrington and McCarthy, 1987; Martin, 2007; Patterson et al., 2007). This scientific humus has provided fertile ground for the rebirth and further development of embodied cognition theories that ground cognition in bodily experience and mental simulation (for a review, see Ghio and Tettamanti, 2010), as opposed to a 'Cartesian' tradition postulating a clear mind/body dualism (e.g. Fodor, 1975). From the point of view of the MNS theory, the finding that macaque mirror neurons apparently code for the core meaning of hand actions, rather than for visual kinematic details – which as a matter of fact may even be obscured with no consequent reduction of mirror neuron activity (Umiltà et al., 2001; Kohler et al., 2002) – has led to a non-associative kind of embodied cognition theories. Such a non-associative semantic MNS theory posits that the human MNS, given its intrinsic perceptuo-motor neurophysiological properties, in addition to action meaning may also mediate the understanding of the semantic meaning of action-related language tokens, such as action verbs (e.g. *to grasp*) or sentences (e.g. *I bite an apple*).

In the experimental practice, associative and non-associative (MNS-based) embodied theories have led to somewhat interchangeable neuroanatomical predictions, but in interpreting the results one has to keep in mind that the underlying mechanistic hypotheses are quite distinct, as highlighted in Section 1. Convergent evidence from behavioral, physiological and neuroimaging studies have demonstrated the involvement of the motor system in processing the action-related semantic content of linguistic utterances. In particular, the processing of verbs referring to actions performed with different body-part (e.g. *to lick*, *to pick*, *to kick*) recruits corresponding somatotopic regions of the motor/premotor cortex (Hauk et al., 2004). Listening to body-part specific action sentences (e.g. *I bite an apple*, *I grasp a knife*, *I kick a ball*) activates a left fronto-parieto-temporal circuit with somatotopic organization in the premotor cortex (Tettamanti et al., 2005). EEG and MEG studies on silent reading of face-, hand-, and leg-related words indicate that the activation of action-related neural networks triggered by the processing of action words is immediate and automatic (Shtyrov et al., 2004; Pulvermueller et al., 2005). For a more comprehensive review of the relevant findings see Pulvermueller and Fadiga (2010), Ghio and Tettamanti (2010).

The finding of somatotopic activations in the left motor and premotor cortices elicited by body-part specific verbs or sentences has been challenged by a number of more recent neuroimaging studies. In an fMRI study, Postle and colleagues (2008) did not find any evidence of activations for body-part specific action verbs matching the somatotopic effects induced by action observation and imitation, but found an action-specific effect common to all effectors in the pre-SMA cognitive motor area. This is compatible with the notion that action verbs entail a more generic form of embodied representation, not directly coupled to motor body schemes. A cross-modal priming study by Galati and colleagues (2008) also failed to show somatotopic activity modulations for hand and mouth actions, when a sound generated by a human action (e.g. hand clapping) was preceded by a congruent masked prime word (e.g. applause) compared to a non-congruent masked prime word (e.g. fly). A possible reconciliation of these divergent findings is offered by a recent proposal by Fernandino and Iacoboni (2010). These authors performed a meta-analysis of the neuroimaging studies on action-related language, showing great neuroanatomical variability between the body-part specific activation foci reported in the different studies (a similar variability was also noted by Postle et al., 2008). They proposed that such a neuroanatomical variability may be explained if we move away from the strict homuncular organization of the motor and premotor cortex known since Penfield and Rasmussen (1950) and instead incorporate more recent findings (Graziano, 2006). These more recent findings suggest that the motor and premotor cortices contain multiple, functionally distinct, representations of each body part, forming a discontinuous topography with many overlapping maps. According to this view, the exact location of the somatotopic foci for action-related language may reflect language-motor correspondences involving different topographic maps that vary according to the type of stimuli (e.g. single words vs. sentences), their linguistic valence (e.g. transitive vs. intransitive), or the task. Such fine-grained, multiple somatotopic representations may, depending on the latter variables, not be undetected, due to the coarse anatomical resolution of fMRI.

In addition to somatotopic representations in the motor/premotor cortex, action-related language also typically engages Broca's area. Broca's area has been proposed to play a role in the integration of compatible and complementary information conveyed by speech and gestures (Gentilucci et al., 2006; Willems et al., 2007), or alternatively to represent actions at an abstract higher-order cognitive level (Tettamanti et al., 2005). More specifically, it is the more anterior pars opercularis (BA 45) which seems to hold more abstract, less context-depending action representations (Nelissen et al., 2005), whereas the more posterior pars opercularis (BA 44) seems to encode for a fine-grained correspondence between executed and perceived motor acts (Gallese et al., 1996). It must be noted, however, that an fMRI study by de Zubicaray and colleagues (2010) failed to provide evidence in favor of the selective involvement of the left inferior frontal cortex (BA44 and BA45) in the comprehension of action verb meanings compared to non-action concrete verbs or non-words. The reasons for this divergent result are unclear, one possibility being again different stimulus selection criteria, such as the use of mixed transitive and intransitive verbs used in isolation, as opposed to full transitive sentences describing the goals and intentions of the agents.

While, overall, this body of evidence points to an existing link between the motor system and semantic language processing, a crucial question is of course whether the motor and premotor cortices actually play a necessary role in the understanding of action-related language. That the role played by the motor and premotor cortices is indeed necessary is more a requisite for non-associative MNS-based than for associative embodied accounts, in which, as described in Section 1, the presence of distributed semantic networks confers a higher degree of information redundancy. Neuropsychological data supporting a necessary causal link have been scarce (e.g. Saygin et al., 2004; Tranel et al., 2003). A more balanced view is therefore to consider the interplay between different cognitive domains, such as language and action, as a dynamic, not an automatic, process, which depends on the specific task at hand for the organism and on the compatibility of the available information (Willems and Haagort, 2007; Mahon and Caramazza, 2008). According to this view, the meaning of action-related language is not coincident with the corresponding motor programs. Rather, the action representation system may mutually interact with perisylvian language areas and other cognitive systems, yielding different nuances of meaning interpretation that vary according to the distributed composition and to the functional and effective connectivity of the overall network (see e.g. Patterson et al., 2007; Ghio and Tettamanti, 2010; Just et al., 2010; Simmons et al., 2010). On the other hand, the involvement of dedicated action-selection neural mechanisms mediated by the left dorsal premotor cortex has been shown to effectively enhance the comprehension of action-related sentences, specifically in sports individuals whose motor skill repertoire reflects the very same actions described by those sentences compared to novice individuals (Beilock et al., 2008). Thus, action-related semantic comprehension is not strictly dependent on the activation of the motor system, but it may be enhanced by its activation.

Finally, we go back to our main argument at stake, namely the perceptual accessibility to the different components of language structure, and, for the case here, to semantics. While the kind of action-related lexical-semantics considered above may in principle be amenable to perceptuo-motor interpretive processes and be successfully tested within the framework of a semantic MNS hypothesis, there are other semantic domains for which this is not the case, such as for words referring to more abstract concepts as the mental/introspective domain (e.g. *to ponder*). Even considering the interesting proposal that part of the semantic representation of abstract concepts may derive from the association with perceptuo-motor metaphors, such as *love-heat* or *grasp the idea* (Gallese and Lakoff, 2005), something for which there is limited and controversial evidence (Aziz-Zadeh et al., 2006; Rueschemeyer et al., 2007; Boulenger et al., 2009), such an account seems unsuited to provide a comprehensive account for all semantic nuances of all abstract concepts (e.g. *guess the opinion*). All in all, a strictly non-associative semantic MNS account based on perceptuo-motor mechanisms appears to be quite limited in scope. Such limitations appear instead to be less of a problem for associative semantic accounts, which can account also for other types of interactions, such as those between perisylvian language parsing regions and not strictly perceptual nor strictly motor systems (Barsalou, 2008; Ghio and Tettamanti, 2010).

## 3. The hidden hierarchy: flattening trees into word strings.

As broadly recognized by virtually all linguistic theories, two major factors are combined when it comes to syntactic structures of all and only human languages (Graffi, 2000). On the one hand, syntactic structures are physically organized as linear sequences of minimal items (morphemes/words), either in a sequence of acoustic elements, of manual and facial configurations of gestures (signs in sign languages), or of written symbols. On the other, syntactic structures are organized hierarchically and this is reflected on the fact that all grammatical functions (such as subject, object, etc.) and complex syntactic relations are defined solely in terms of hierarchical relationships, completely disregarding the linear order. We will illustrate this crucial point by reproducing two prototypical empirical cases involving two independent syntactic processes: pronoun interpretation and question formation.

Pronouns such as *he* can either refer to another noun (phrase) or have independent reference, for example: *John thinks he is smart*, where *he* can either refer to *John* or to another individual in the discourse. As a matter of fact, the capacity to refer or not to refer to other elements is highly restricted by syntax. For example, if we simply invert the order of *John* and *he* in the sentence given here we get the following sharp contrast: *He thinks that John is smart*, where *he* can no longer refer to *John*. One intuitive explanation for this, albeit false, would be to rely on the linear order of words and say that a pronoun, given that it replaces a noun phrase, can not precede the noun it refers to. In fact, this solution can be easily disproved by considering cases like the following: *When Mary said he was always late, John blamed it on the Italian railway system*, where *he* can refer to *John* even if it precedes it. A canonical way to account for this phenomenon is to rely on the notion of syntactic prominence, which is a syntactic relationship based on two-dimensional (i.e. non-linear) structural boundaries and constraints, usually expressed in terms of syntactic trees. In fact, the generalization that emerges on the base of comparative research is that a pronoun like *he*, and its equivalent across languages, can not refer to any other Noun Phrases in the smallest clause structure it is contained in (technically, it must be “free”), disregarding the linear order between the pronoun and the noun phrase.

Another case study, supporting the same conclusion concerning the dependence of complex syntactic relations on hierarchy rather than on linear order, comes from those relations that *prima facie* only affect the word order of affirmative active sentences without introducing any new morpheme. A simple sentence like *John is smart* can be turned into an interrogative matrix sentence by changing the order of the copula with respect to the subject, yielding *Is John smart?*. When two tokens of the copula occur in the same sentence it can be easily shown that linear order does not count at all for this type of operations. So, for example, starting from *John is interested in knowing if Mary is blonde*, we only get *Is John interested in knowing if Mary is blonde?* as a grammatical sentence, as opposed to **Is John is interested in knowing if Mary blonde?*. Of course, this can not be captured by simply assuming that only the first occurrence of the copula is visible to the word order rearranging operation (movement), witness cases like the following: *John who is interested in Mary is buying a new car*, where only the second occurrence of the copula can be moved, making word-order-based operations untenable: *Is John who is interested in Mary buying a new car?* vs. **Is John who interested in Mary is buying a new car?* The intuition here is that only the copula contained in the matrix (i.e. main) clause structure is visible to movement and since the very notion of matrix clause is a hierarchical notion, we are forced to conclude that it is hierarchy that matters in movement operations. Similar arguments would support the extension of this conclusion for other syntactic phenomena (see Rizzi, 2009), including locality (i.e. those principles that limit structural dependencies (Manzini, 1992; Carnie, 2002)) and morphosyntactic phenomena (such as prototypical agreement facts).

Formally speaking, a distinguishing feature of syntactic hierarchies, as opposed to hierarchical structures in other grammatical domains (e.g. syllable structure), is that they can be generated by recursion.[2] Recursion as found in human language syntax is one that implies embedding of a certain structure within the same structure, something which is common in all languages of the world and that has been first captured formally since the seminal works by Chomsky in the 1950s. Chomsky (1957) provided a formal proof that the application of stochastic algorithms such as markovian chain models to natural language syntax, as suggested by Shannon (1949) among others, was insufficient to account for syntax: in particular, any language contains a potentially infinite set of non-local dependencies, whereby the occurrence of a certain word instantiates a dependency with another word in the same sentence at an unbounded distance. A simple case is the well-known example of the dependence between *if* and *then* that could be separated by as many words as the memory capacity of the speaker allows, but any language can contain a potentially infinite set of non-local dependencies depending on the lexicon it is based on.

This kind of reasoning has been captured formally by means of recursion theory (for an introduction, see Hopcroft et al. (2007)). In fact, the proof that natural languages can not be captured by stochastic models based on Markovian chains was proved in two steps by Chomsky (1957). First, consider a simple model consisting of a language containing only two words (a and b) and a set of sentences where a certain number of a is followed by the same number of b (call it $a^nb^n$). Crucially, it was shown that this language can be generated only by a context free grammar, that is a phrase structure rule of the type Z → aZb | ab, where the non-terminal symbol Z may be rewritten either as Z or as aZb. Second, it was shown that any natural language contains analogous cases of dependencies. The combination of these two results provided the proof that that stochastic models based on Markovian chains are inadequate to capture locality principles of natural languages even at a mere descriptive level. There are of course other kind of grammars

[2] Syllables are hierarchical, in that they contain an onset and a rhyme and the rhyme contains the nucleus and the coda on its turn, but they can not be generated by recursion, because neither the onset nor the ryme can contain another onset or a rhyme, respectively.

that can not be generated by stochastic models, which are simple cases of languages, including the so called ww and ww$^r$, where a certain sequence of symbols is followed by the same sequence or by its mirror image, respectively (see again Hopcroft et al. (2007) for a general treatise of this issue). Interestingly, this result has been taken into account by all models of generative grammars, including both transformational and non-transformational approaches (for a historical overview, see Graffi (2000)).

With respect to a specific class of non-transformational generative approaches, i.e. usage-based approaches such as Construction Grammar (Goldberg, 2006) or Head-Driven Phrase Structure Grammar (Pollard and Sag, 1994), some further considerations may be relevant. According to usage-based generative approaches, children in the early stages of development do not acquire language on the basis of innate grammatical categories and syntactic rules, but rather gradually by mastering an increasing set of lexically specific constructions that can range from single words to more abstract constructions. However, note that the proposed usage-based grammars are formally equivalent to the context free grammar proposed by Chomsky (1957), including the non-linear properties of non-local hierarchical dependencies detailed above. Thus, even though usage-based approaches pose the emphasis on the accumulation of specific lexical knowledge driving early linguistic development, rather than on rule-based syntactic structuring of input lexical entries, the same caveats regarding the unattainability of non-linear hierarchical syntactic properties to the bodily senses still apply. More specifically, it is not so that usage-based accounts, as opposed to transformational accounts, can be advocated to support a MNS-based theory of syntactic processing. Indeed, a lot more of linguistic structures than previously thought can probably be learned by children in the early developmental stages by relying on probabilistic inferences about multiword combinations, as demonstrated by a Bayesian computational procedure (Bannard et al., 2009); but with respect to later developmental stages, the efficacy of these procedures is sub-optimal if they do not incorporate a restricted model of grammatical abstraction and generalization (Bannard et al., 2009). Thus, the implications of usage-based approaches for a perceptuo-motor theory of syntax are largely equivalent to those of any other generative grammar approach: namely, that MNS can mediate those aspect of language processing that are readily available through hearing or vision, but the progress toward a full-fledged linguistic capacity can most likely only be achieved through the maturation of non-perceptual, internal cognitive skills related to hierarchical structuring at later developmental stages.

In this respect, syntax may interact with the MNS in similar proportions as phonology and semantics, i.e. some aspects of syntax may be amenable to perceptuo-motor mapping processes, but some are not (and in the case of syntax, indeed some very crucial and pervasive aspects, as we have seen). Some syntactic phenomena are presumably transparent in terms of linear sequential information, such as for instance simple multiword binding between nouns and verbs, between simple noun and verb phrases, or idiomatic expressions such as “kick the bucket”. But even then, it remains to be demonstrated that all or even just some of these phenomena indeed depend on the activity of the visuo-motor MNS.

The available neuropsychological and neuroimaging evidence on the neural correlates of processing syntactic hierarchies (Embick et al., 2000; Moro et al., 2001; Tettamanti et al. 2002, 2009; Musso et al., 2003; Friederici et al., 2006; Bahlmann et al., 2008; Makuuchi et al., 2009; Shetreet et al., 2009), as well as other non-local syntactic phenomena, such as syntactic movement (Caplan, 2001; Ben-Shachar et al., 2004; Grodzinski, 2006; Santi and Grodzinski, 2010) or argument hierarchies (Bornkessel et al., 2005), consistently demonstrate a crucial role of BA 44 in the left hemisphere. Although this body of evidence needs to be considered carefully – due to the limited spatial and temporal resolution of the employed anatomo-function mapping methods with respect to single cell neurophysiological recordings of mirror neuron activity, and also due to the lack of specifically aimed research questions – it seems that beyond of the involvement of Broca's area, the processing of syntactic hierarchies does not consistently involve other core MNS regions, such as the premotor cortex or the rostral inferior parietal cortex, as should be expected if the MNS is to play a relevant role in the perceptuo-motor mapping of non-local syntactic dependencies. One of these core MNS, regions, namely the ventral premotor cortex, has been shown by one fMRI study to be specifically engaged in the processing of local syntactic dependencies, as opposed to Broca's are for non-local dependencies (Opitz and Friederici, 2007). This finding is interesting if one consider, as we do propose, that local, but not non-local, syntactic dependencies may in certain languages be transparent to the bodily senses, as long as they involve agreement morphemes that are detectable by the ear.

Summarizing, despite the fact that grammar is physically encoded as a linear code, the underlying syntactic structure is hierarchically organized. As a consequence of linearisation, the hierarchical structure underlying the bi-dimensional appearance of sentences is not entirely transparent to the bodily senses (for a detailed discussion of linearisation based on a formal model, see Kayne (1994) and Moro (2000, 2008))[3]. Thus, a perceptuo-motor theory of syntax comparable to the motor theory of speech perception does not appear to be viable.

Importantly, we do not claim that syntactic abstract rules are not embodied in the

[3] The central notion of Kayne's (1994) antisymmetry theory, namely the linear correspondence axiom (LCA), must not be interpreted as implying that hierarchy can be inferred from the linear order. There are at least two reasons that makes this conclusion untenable. One is assumed by definition: the empirical content of the LCA is in fact that there is only one linear ordering (not to be confused and identified solely with the linear sequence of words) and that this is manifested both as the linear sequence of words (which we can of course observe) and the suitable hierarchical structure (which we can not). The other is empirical. Given that languages involve movement and non-phonological categories, it is impossible to derive the hierarchical structure underlying any given linear ordering in an unambiguous way, weather or not the LCA is adopted. Any element in the sequence could be potentially moved from an underlying position adding more structure, or there could simply be other syntactic categories that are not pronounced but that are there in the sequence. If it were not so, the LCA would just turn out to be at best a tautological statement on language structure, and a rather trivial one. The same conclusion holds for the weak version of the antisymmetry theory, Dynamic Antisymmetry, proposed in Moro (2000).

brain: they must be, and progress in neuroscience will tell how and what kind of neural mechanisms are necessary for syntactic processing, as also pointed out by Pulvermueller (2010). Rather, given currently available evidence, we argue against the view that syntax, and particularly hierarchical structuring, is embodied in the MNS, as has been suggested for instance by Pulvermueller and Fadiga (2010). In a recent review article on the role of the sensorimotor system in language processing, they discuss neuropsychological and neuroimaging evidence on the involvement of the pars opercularis of Broca's area in syntactic processing and in sequential processing in other cognitive domains as favoring the view that the action-perception circuit mediates grammar processing, possibly including embedded hierarchical syntactic structures. But to show that the polymodal opercular portion of Broca's area is involved, at the macro-anatomical level, in both syntactic and action processing is not sufficient to demonstrate the direct role of mirror neurons and of the MNS in both domains, particularly given that the MNS is thought to extend to other brain regions not typically involved in syntactic processing, such as the premotor and the rostral inferior parietal cortex. Pulvermueller and Fadiga (2010) seem to be well aware of these caveats, as they do never make explicit reference to a putative direct involvement of mirror neurons in syntactic processig. The same lack of explicit reference to the MNS applies to the discussion of a neuronal model capable to handle embedded hierarchical structures (see also Pulvermueller and Knoblauch (2009) and Pulvermueller (2010))[4].

In our view, claims such as the one by Pulvermueller and Fadiga (2010) are also problematic, due to some rather common terminological and conceptual confusions regarding the hierarchical properties of the language syntax, as we will now detail in the

[4] Incidentally, we note that constant progress is being made in developing computational models of syntactic processing (e.g. Crocker, 1999; Pulvermueller and Knoblauch, 2009), including model-based learning of embedded hierarchical structures (but admittedly still failing to account for other, more complex syntactic phenomena, such as locality discussed above). Hopefully, this development will eventually provide us with a more thorough understanding of the embodied neuronal mechanisms underlying syntactic processing. However, in our view, the properties of such models make them more suitable to account for internal computational mechanisms than for perceptuo-motor mapping mechanisms.

following paragraphs.

The failure to fully appreciate the implications of how linear compression affects human language syntax leads to fallacious shortcomings in the assimilation of the language domain to other cognitive domains, an issue that we have already discussed elsewhere (Tettamanti and Weniger, 2006). For instance, in the fMRI study by de Zubicaray and colleagues (2010) mentioned in the previous section, the lack of action word specific semantic effects is taken to support the hypothesis of a role of Broca's area in the “hierarchical linearisation” of information across action and linguistic domains. In fact, the term here is used in a rather misleading metaphorical way as opposed to the formal sense that it has acquired in theoretical linguistics. More specifically, their experimental paradigm did not include an explicit manipulation of hierarchical vs. linear structures and they did not provide any formal definitions of “hierarchical linearisation” in each of the two domains. This problem has been recognized by Rizzolatti and Craighero (2004), who noted that there is no experimental evidence in support of the speculation that the pars triangularis of Broca's area may code the alleged “syntactic” aspects of action. As another example of this confusion – recalling the discussion about the motor system and language of the preceding section – Glenberg (2007) conceived a sort of proto-syntax, in which thoughts, as they are expressed by language, are combined by mechanisms of action control. The main problem with these types of motor syntactic accounts is that there is an apparent confusion between the perceptual level, where a link between language and the motor systems has been established, with a non-perceptual or internal level. We refer to an internal level as a set of cognitive or computational skills that can be used to process, integrate or transform neural information, such as the processing of the hierarchically organized structures found in human language syntax. It may well be the case that primary speech-related information is gained by the perceptual systems, but the intermediate and final products of syntactic processing are in most instances higher order association products.

The internal computational skills that instantiate hierarchical relations and recursion may in principle operate supramodally across different cognitive domains, within different contexts and with different degrees of complexity (Tettamanti and Weniger, 2006). For example, recursive hierarchical structuring has indeed been implicated, in addition to language, in action control (Greenfield, 1991), visuo-spatial functions (Tettamanti et al., 2009; Bahlmann et al., 2009), spatial navigation (Hauser et al., 2002), and music (Patel, 2003). But precisely because these mechanisms are internal and not dependent upon perception, we can not assume that their neural basis is constituted by mirror neurons or by equivalent perceptuo-motor matching systems[5].

Shortcomings in the assimilation of the language domain to other cognitive domains may also derive from confusion regarding a different use of the term hierarchical organization in cognitive psychology. In this field, this term has often been used to refer to a quite different phenomenon than the one we alluded to above with respect to syntax, namely sequential hierarchy. Lashley (1951), and more recently Koechlin and Jubault (2006) among others, have employed the notion of sequential hierarchical organization in developing accounts of action execution and action planning. There, for instance, the action of reaching and grasping an object with the hand is encoded in a motor plan with a precisely defined sequence of events that are hierarchically ordered, such as for example: moving the forelimb first, then opening the fingers, and finally grasping and lifting the object. This principle may also apply to the planning of facts, events or decisions, such as for example: the action to be executed in response to hearing a telephone ring may be

[5] It may be of some interest, though, that a connectionist neural network model has been proposed, in which the representation of phrase constructions is implemented, by some degrees of analogy to the MNS, as goal states that can be recognized both from produced and perceived symbol strings (Womble and Wermter, 2001). This connectionist model has been shown to learn and generalize context free syntactic rules generated by recursion. Even if we leave aside the fundamental issue of locality restrictions, such as those exemplified above, which are not accounted for by this model, the enthusiasm for such an account of the acquisition of syntactic competence and syntactic processing is undermined by another caveat. The MNS module as implemented within the connectionist model is non-specific, since the representation of nested hierarchies is delegated to memory stacks with unspecified neurophysiological properties. This actually reinforces our argument of the unavailability of syntax to the sensorimotor system.

hierarchically controlled by contextual information (e.g. the subject is or is not the owner of the ringing telephone) or by specific goals (e.g. the subject pretends not to be at home). Or, as in another example proposed by Pulvermueller and Fadiga (2010), the act of embedding an action, such as turning off the light, between an independent sequence of related actions, such as door opening and door closing. This kind of hierarchically organized sequential information may be perceived by the bodily senses, and the MNS may in principle be involved in the perceptuo-motor transformation of some of these action sequences (Schubotz and Von Cramon, 2001, 2002a, 2002b), but such sequential processes are not a relevant support for a perceptuo-motor theory of human language syntax, as long as it is based on hierarchical relations that are not visible in the linear sequence of the linguistic physical code (as noted above, some relations in certain languages may be signalled by agreement phenomena).

This sort of terminological and conceptual confusion between different meanings of hierarchical organization (internal vs sequential) emerges also in proposals on the common evolutionary origin of language and motor system and of Broca's areas' functions. For instance, Fazio and colleagues (2009) suggested that Broca's area might have specialized both in hierarchical syntactic computations and in a sort of motor syntax that specifies the rule-based hierarchical composition of simple motor acts into actions with a goal. Although in principle possible, what these theories leave fully unspecified is the mechanism by which the MNS, as the underlying shared neuronal substrate, may have evolved its function from the more rudimentary capacity of computing *sequential* hierarchical information (in primates) to the capacity of computing *non-linear* hierarchical relations as found in language syntax (in humans). Furthermore, leaving the evolutionary problems aside, these theories do not specify how, in humans, a particular class of neurons with perceptuo-motor properties become active in response to both sequential and non-linear hierarchical information, whereby the former information is available to the senses and the latter not. Without a clear account of these two problems, such hypotheses can not be translated into testable experimental predictions.

Thus, we can conclude that neither the processing nor the acquisition of syntax (both ontogenetically and phylogenetically) can (so far) be traced back to the activity of the MNS. On the other hand, as we will argue in the next section, it is reasonable to assume that the non-perceptual, computational system underlying syntax and the perceptuo-motor system, though independent, closely interact in language processing (see Arbib (2010), for similar conclusions but from a very different neurolinguistic perspective): we will provide such a case by illustrating some aspects of neurophysiological phenomena related to sentential negation.

## 4. **When syntax modulates action representation**

Let us first recapitulate the main argument of MNS theories of language processing. The phonological and semantic components of language may have evolved from the capacity, already present in non-human primates, to understand actions performed by others. This capacity may have been empowered in humans with a greater ability to imitate and with a broader spectrum of neurophysiological responses of the MNS, including pantomimes and intransitive actions. As a consequence, the ability of the MNS to understand the meaning of manual gestures may have been extended to also include oro-laryngeal gestures (Rizzolatti and Craighero, 2004). This hypothesis has been invoked to predict that the same brain areas are activated when executing, observing, imitating a motor gesture, and when hearing a linguistic token produced by or describing the same gesture. Of course, this is not equivalent to assuming that the language faculty as a whole has evolved from the MNS. What appears as a more conservative view, in the light of the many caveats reviewed above, is that *some* aspects of semantic and phonetic processing related to perceptuo-motor functions, and possibly also the capacity to establish some simple syntactic combinations between multiple lexical entries, may have evolved in connection with the MNS. These perceptuo-motor associated components of the language grammar may have interacted and may have been integrated with a set of internal computational capacities, i.e. syntactic hierarchical structuring and recursion, and possibly with other cognitive capacities, to give rise to the full-fledged language faculty.

Is there any available evidence that such an interaction between perceptuo-motor associated language components and syntactic computations may actually occur? Based on recent findings in our laboratory, we propose that a case study for this kind of interactions is provided by sentential negation. From a semantic point of view, sentential negation is a logical operator that inverts the truth value of a given affirmation, so that *John has killed a snake* and *John has not killed a snake* have opposite conditions for their truth value to be assigned. From a morphosyntactic point of view, sentential negation has interesting properties. First of all, unlike say interrogative sentences, to express a negative sentence a specific morpheme is involved in all languages of the world (Horn, 1989; Zanuttini, 1997). More explicitly, there are no languages where negation can be created by simply changing the word order of an affirmative sentence, as opposed to say interrogatives (cf. *John is smart* vs. *Is John smart?*). This is already an interesting sign that negation can not be a “purely” syntactic phenomenon. Furthermore, negation can be applied recursively, so that if we negate a sentence like *John thinks that you can’t make it* turning it into *John doesn't think that you can’t make it*, we get a sort of neutralization of the two negations, yielding something like *John thinks that you can make it*: This points to the fact that the interpretation of negation, according to what we have discussed in the previous section, comes from the application of recursive rules, although it does not reduce to the application of recursive rules, for all structures must be uttered appropriately, that is in pragmatically felicitous way in the sense of Grice (1989) according to the context.

Notice also that the interpretation of negation may very well be considered as a paradigmatic case in that it shows that it is hierarchical structures that matter rather than linear order in syntax. The following simple case illustrates the point. Consider the sentence *I do <u>not</u> know why John has left <u>nor</u> why*. In this sentence the negative subordinative conjunction *nor* must be preceded by negation but precedence is not sufficient. This can be seen by considering the contrast with another sentence: **I know why John has <u>not</u> left <u>nor</u> why*. The ungrammaticality of the second sentence can be easily explained as follows, by assuming that for nor to be interpreted it must occur within the scope of negation, that is within the set of nodes contained in the node which is adjacent to negation. By using brackets to identify the scope of negation in the two sentences here, we get: *I do [<u>not</u> [know [why John has left] <u>nor</u> [why]]]* vs. *I know why John has [[ <u>not</u> [left]] <u>nor</u> [why]]*. Clearly, *nor* is not in the scope of *not* in the second sentence, even if *not* precedes *nor* exactly as in the first sentences, thus proving that hierarchical structures is what explains the interpretation of negation as far as syntax is concerned (for a detailed analysis of this phenomena see Moro (in press) and references cited there).

Finally, a futher important aspect in relation to our argument of the unattainability of syntax to the bodily senses is that the information carried by sentential negation does not correspond to any event that can be perceived in the physical world. More explicitly, there is nothing in the world that embodies negation, so to speak: the world can only be made of positive facts, i.e. affirmations. For instance, if we say *The cake is not on the table*, there is no sensory input that can directly elicit the content of this sentence, unless one would say that perceptual input can derive from a potentially infinite number of other facts that involve a cake which is not “being on a table”. If I see a table empty, that can by no means inform me that there is not a cake on the table, because there are infinite stentences about negative facts that would be derivable from the same context, such as *The car is not on the table*. In turn, when I actually see a cake lying on a table, the sensory information about the state of the cake is univocal. Hence, the meaning of *The cake is not on the table* can not be entirely derived from perceptual stimuli. Contextual and perceptual information may support or even speed up the interpretation of negative sentences, as suggested by the fact that in naturalistic settings negative sentences are typically not uttered unless the proposition being negated was previously explicitly mentioned (Kaup et al., 2007). But even given the prior contextual availability of pragmatic cues, such as the binary choice *The cake is either on the table or in the fridge*, the neural mechanisms responsible for (negative) linguistic comprehension must be flexible enough to cope with an indefinite number of alternative solutions: even if, following this binary choice, we say *The cake is not on the table*, this does not necessarily imply that the cake is in the fridge, because it may have been already eaten up. In other words, the interpretation of sentential negation can not be reduced to any property referring to the ease of comprehension, to the context, to pragmatical rules or, crucially, to any sensorimotor stimulus.

In sum, sentential negation is a syntactic recursive operation that is per se unrelated to any event that may be perceived by the bodily senses. But when sentential negation is applied to reverse the truth value of action-related sentences, a unique type of interaction with the perceptuo-motor system is observed. An fMRI study by Tettamanti and colleagues (2008) demonstrated that the kind of involvement of the motor system that, as we have described above, occurs during the processing of affirmative action-related sentences such as *I bite an apple*, is significantly reduced in the case of corresponding negative sentences such as *I do not bite an apple*. Reduced activations for negative compared to affirmative action-related sentences were observed in Broca's area and in left hemispheric fronto-parieto-temporal regions, including the premotor cortex, the rostral inferior parietal cortex, and the posterior temporal cortex; that is to say, in all main brain regions constituting the MNS. This kind of interaction between a syntactic operation (sentential negation) and semantic action-related representations was interpreted as a neurophysiological correlate of a transiently reduced access to the semantic information affected by negation, that is the predicate expressed by the verb, compatible with psycholinguistic accounts (Kaup et al., 2007). In other words, the perceptuo-motor mapping elicited by affirmative action-related sentences, which, as we have seen, may enhance their semantic interpretation possibly through mental simulation, is at least transiently blocked in the case of negative sentences. Furthermore, using an effective connectivity analysis, we also tested how the functional cross-talk between perisylvian language parsing regions and action-related representations is modulated by sentential negation. Compared to affirmative action-related sentences, the modulatory connection strengths from the left premotor, rostral inferior parietal and posterior temporal cortices to Broca's area were significantly reduced during the processing of corresponding negative sentences. As far as the general framework of associative embodied language theories is concerned, this suggests that under the scope of sentential negation, the action-related semantic information represented in perceptuo-motor systems is integrated to a lesser extent into higher-order linguistic representations and possibly into non-perceptual inferences of the kind illustrated in the previous paragraph[6].

Of course, many different steps can be taken to (dis)prove this theory. In particular, one could see whether such interactions between syntax and semantic representations occur mutually in both directions – as expected in a theory of independent neural systems that flexibly cooperate to form distributed associative networks – or whether they are instantiated asymmetrically. One possible line of research we are pursuing is to combine experiments on the influence of sentential negation on action kinematics with symmetrical experiments on the influence of motor gestures on the processing of sentential negation.

[6] Note that we can safely exclude what may constitute a more parsimonious explanation of the findings by Tettamanti et al. (2008), an explanation that would contradict the main tenet of the present article, namely: it may be concluded that the modulation of the MNS induced by negative action-related sentences does not reflect a syntactically driven reduction of the access to semantic information, but more simply the very syntactic processing of sentential negation. If so, this would constitute a strong evidence in favor of the view that the MNS can indeed mediate syntactic processing. However, the same patterns of reduction of activation and of effective connectivity in the MNS were not observed for sentences with an abstract content (e.g. “I (do not) appreciate loyalty”). In this latter case, a reduction of activation and of effective connectivity was found in a left temporal and medial network specific for the processing of abstract sentences, pointing to a modulation of semantic processing rather than syntactic processing per se (Tettamanti et al., 2008).

Preliminary findings from the first of these two symmetric experiments indicate, in agreement with the fMRI results (Tettamanti et al., 2008), that during the processing of negative vs. affirmative action-related sentences the perceptuo-motor system is under reduced semantic processing demands, leaving more computational resources available to support the planning and execution of a motor gesture (Bartoli et al., 2010).

## 5. Conclusions

We have provided what in our views are the most relevant empirical and theoretical arguments against a comprehensive MNS based theory of language processing. In particular, we argued against the view that the MNS encodes perceptuo-motor mechanisms that may directly capture the non-linear hierarchical structure of the human language syntax in the same way as they may capture action sequences. More specifically, we suggested that the linearisation of hierarchical structures makes the relevant syntactic information unavailable to sensory perception. Rather, human language is characterized by a highly intricate structure at all levels of the grammar, and each of these levels, including phonetics and lexical-semantics, requires the integration of internal computational mechanisms and perceptuo-motor systems. This kind of integration is particularly interesting in the case of sentential negation, in which syntactic computational mechanisms, which can not be traced back to any perceptual stimulus, and the perceptuo-motor system closely interact to map the intended semantic meaning. Thus, perceptuo-motor mapping mechanisms may well contribute to higher-order associative neural networks that instantiate the uniquely human language faculty, provided that we keep their crucial role distinct.

## Acknowledgments

We are grateful to Marta Ghio, Valentina Bambini, and to three anonymous reviewers for helpful comments on our manuscript.

**References**

Arbib MA. Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain and Language*, 112 (1):12-24, 2010.

Aziz-Zadeh L, Wilson S, Rizzolatti G, and Iacoboni M. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16:1818-1823, 2006.

Bahlmann J, Schubotz RI, and Friederici A. Hierarchical artificial grammar processing engages Broca's area. *Neuroimage*, 42 (2):525–534, 2008.

Bahlmann J, Schubotz RI, Mueller JL, Koester D, and Friederici A. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*, 1298:161-170, 2009.

Bannard C, Lieven E, and Tomasello M. Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences, U.S.A.*, 106 (41):17284-17289, 2009.

Barsalou L. Grounded cognition. *Annual Review of Psychology*, 59:617-645, 2008.

Bartoli E, Caporizzo A, Tettamanti A, Gatti R, Moro A, Perani D, and Tettamanti M. Better forelimb motor performance when listening to negative vs affirmative action-related sentences. Abstract presented at the *Workshop on Concepts, Actions, and Objects: Functional and Neural Perspectives. 4th Edition*. Rovereto, Italy, 2010.

Beilock SL, Lyons IM, Mattarella-Micke A, Nusbaum HC, and Small SL. Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences, U.S.A.*, 105 (36):13269-13273, 2008.

Ben-Shachar M, Palti D, and Grodzinsky Y. Neural correlates of syntactic movement: converging evidence from two fMRI experiments. *Neuroimage*, 21 (4):1320-1336, 2004.

Bornkessel I, Zysset S, Friederici AD, von Cramon DY, and Schlesewsky M. Who did what to whom? The neural basis of argument hierarchies during language comprehension. *Neuroimage*, 26 (1):221-233, 2005.

Boulenger V, Hauk O, and Pulvermüller F. Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral Cortex*, 19 (8):1905-1914, 2009.

Buccino G, Binkofski F, Fink G, Fadiga L, Fogassi L, Gallese V, Seitz R, Zilles K, Rizzolatti G, and Freund H. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13:400-404, 2001.

Caplan D. Functional neuroimaging studies of syntactic processing. *Journal of Psycholinguistic Research*, 30 (3):297-320, 2001.

Carnie A. *Syntax*. Oxford, UK: Wiley-Blackwell, 2002.

Chomsky N. *Syntactic Structures*. The Hague: Mouton, 1957.

Chomsky N. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

Corballis MC. Mirror neurons and the evolution of language. *Brain and Language*, 112 (1):25-35, 2010.

Crocker MW. Mechanisms for sentence processing. In Garrod SC, and Pickering MJ (Eds), *Language processing*, London, UK: Psychology Press, 1999.

D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, and Fadiga L. The motor somatotopy of speech perception. *Current Biology*, 19 (5):381-385, 2009.

Embick D, Marantz A, Miyashita Y, O Neil W, and Sakai K. A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences, U.S.A.*, 97 (11):6150-6154, 2000.

Fadiga L, Craighero L, Buccino G, and Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15 (2):399-402, 2002.

Fazio P, Cantagallo A, Craighero L, D'Ausilio A, Roy AC, Pozzo T, Calzolari F, Granieri E, and Fadiga L. Encoding of human action in Broca's area. *Brain*, 132 (Pt 7):1980-1988, 2009.

Fernandino L and Iacoboni M. Are cortical motor maps based on body parts or coordinated actions? Implications for embodied semantics. *Brain and Language*, 112 (1):44-53, 2010.

Fischer MH and Zwaan RA. Embodied language: a review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, 61 (6):825-850, 2008.

Fodor JA. *The language of thought*. Cambridge, MA: Harvard University Press, 1975.

Friederici A, Bahlmann J, Heim S, Schubotz RI, and Anwander A. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences, U.S.A.*, 103 (7):2458–2463, 2006.

Galantucci B, Fowler CA, and Turvey MT. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13 (3):361-377, 2006.

Galati G, Committeri G, Spitoni G, Aprile T, Di Russo F, Pitzalis S, and Pizzamiglio L. A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage*, 40 (3):1274-1286, 2008.

Gallese V, Fadiga L, Fogassi L, and Rizzolatti G. Action recognition in the premotor cortex. *Brain*, 119 (2):593-609, 1996.

Gallese V, and Lakoff G. The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22:455-479, 2005.

Gentilucci M, Bernardis P, Crisi G, and Dalla Volta R. Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. *Journal of Cognitive Neuroscience*, 18 (7):1059-74, 2006.

Ghio M and Tettamanti M. Semantic domain-specific functional integration for action-related vs. abstract concepts. *Brain and Language*, 112 (3):223-232, 2010.

Glenberg A. Language and action: Creating sensible combinations of ideas. In Gaskell M (Ed), *The Oxford handbook of psycholinguistics*, Oxford, UK: Oxford University Press, 2007.

Goldberg AE. *Constructions at work: the nature of generalization in language*. Oxford, UK: Oxford University Press, 2006.

Golfinopoulos E, Tourville JA, and Guenther FH. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, 52 (3):862-874, 2010.

Graffi G. *200 years of syntax*. Amsterdam: John Benjamins Publishing Company, 2000.

Graziano M. The organization of behavioral repertoire in motor cortex. *Annual Review of Neuroscience*, 29:105-134, 2006.

Greenfield P. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, 14:531-595, 1991.

Grice P. *Studies in the way of words*. Cambridge, MA: Harvard University Press, 1989.

Grodzinsky Y. The language faculty, Broca's region, and the mirror system. *Cortex*, 42 (4):464-468, 2006.

Guenther FH, Ghosh SS, and Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96 (3):280-301, 2006.

Hauk O, Johnsrude I, and Pulvermuller F. Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41:301-307, 2004.

Hauser M, Chomsky N, and Fitch W. The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298:1569-1579, 2002.

Hickok G. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21 (7):1229-1243, 2009.

Tettamanti M. & Moro A. *Cortex*, 48(7), 923–935 (2012).

Hopcroft JE, Motwani R, and Ullman JD. Introduction to automata theory, languages, and computation. Atlanta, GA: Pearson/Addison Wesley, 2007.

Horn L. *A Natural History of Negation*. Chicago: University of Chicago Press, 1989.

Jackendoff R. *Foundations of language*. Oxford, UK: Oxford University Press, 2003.

Just MA, Cherkassky VL, Aryal S, and Mitchell TM. A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PloS One*, 5 (1):e8622, 2010.

Kaup B, Zwaan R, and Lüdtke J. The experiential view of language comprehension: How is negation represented? In Schmalhofer F, and Perfetti C (Eds), *Higher level language processes in the brain: Inference and comprehension processes*, Mahwah, NJ: Erlbaum, 2007.

Kayne R. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press, 1994.

Koechlin E and Jubault T. Broca's area and the hierarchical organization of human behavior. *Neuron*, 50 (6):963-974, 2006.

Kohler E, Keysers C, Umilta M, Fogassi L, Gallese V, and Rizzolatti G. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297 (5582):846-848, 2002.

Kotz SA, D'Ausilio A, Raettig T, Begliomini C, Craighero L, Fabbri-Destro M, Zingales C, Haggard P, and Fadiga L. Lexicality drives audio-motor transformations in Broca's area. *Brain and Language*, 112 (1):3-11, 2010.

Lashley K. The problem of serial order in behavior. In Jeffress L (Ed), *Cerebral mechanisms in behavior: The Hixon Symposium*, New York: Wiley, 1951: 112-146.

Liberman AM and Mattingly IG. The motor theory of speech perception revised. *Cognition*, 21 (1):1-36, 1985.

Lotto AJ, Hickok GS, and Holt LL. Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13 (3):110-114, 2009.

Tettamanti M. & Moro A. *Cortex*, 48(7), 923–935 (2012).

Mahon BZ and Caramazza A. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, Paris*, 102 (1-3):59-70, 2008.

Makuuchi M, Bahlmann J, Anwander A, and Friederici A. Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences, U.S.A.*, 106 (20):8362-8367, 2009.

Manzini MR. *Locality: A Theory and Some of Its Empirical Consequences*. Cambridge, MA: MIT Press, 1992.

Martin A. The representation of object concepts in the brain. *Annual Review of Psychology*, 58:25-45, 2007.

Moro A. *Dynamic Antisymmetry*. Cambridge, MA: MIT Press, 2000.

Moro A. *The boundaries of Babel*. Cambridge, MA: MIT Press, 2008.

Moro A. Clause structure folding and the wh-in situ effect. *Linguistic Inquiry*, in press.

Moro A, Tettamanti M, Perani D, Donati C, Cappa S, and Fazio F. Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage*, 13 (1):110-118, 2001.

Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, Buchel C, and Weiller C. Broca's area and the language instinct. *Nature Neuroscience*, 6 (7MMG+03):774-781, 2003.

Nelissen K, Luppino G, Vanduffel W, Rizzolatti G, and Orban GA. Observing others: multiple action representation in the frontal lobe. *Science*, 310 (5746):332-336, 2005.

Opitz B, and Friederici A. Neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping*, 28 (7):585-592, 2007.

Patel A. Language, music, syntax and the brain. *Nature Neuroscience*, 6 (7):674-681, 2003.

Patterson K, Nestor P, and Rogers T. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Review*

*Neuroscience*, 8 (12):976-87, 2007.

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, and Rizzolatti G. Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91 (1):176-180, 1992.

Penfield W and Rasmussen T. *The cerebral cortex of man*. New York: Macmillan, 1950.

Pollard C, and Sag I. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press, 1994.

Postle N, McMahon KL, Ashton R, Meredith M, and de Zubicaray GI. Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *Neuroimage*, 43 (3):634-644, 2008.

Pulvermuller F. Brain embodiment of category-specific semantic memory circuits. In Semin GR, and Smith ER (Eds), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*, Cambridge, UK: Cambridge University Press, 2008: 71-97.

Pulvermüller F. Brain embodiment of syntax and grammar: discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain and Language*, 112 (3):167-179, 2010.

Pulvermuller F, Shtyrov Y, and Ilmoniemi R. Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, 17 (6):884-92, 2005.

Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, and Shtyrov Y. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences, U.S.A.*, 103 (20):7865-7870, 2006.

Pulvermüller F, and Knoblauch A. Discrete combinatorial circuits emerging in neural networks: a mechanism for rules of grammar in the human brain? *Neural Networks*, 22 (2):161-172, 2009.

Pulvermüller F, and Fadiga L. Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11 (5):351-360, 2010.

Rizzi L. The discovery of language invariance and variation, and its relevance for the cognitive sciences. *Behavioral and Brain Sciences*, 32:467-468, 2009.

Rizzolatti G, Fogassi L, and Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2 (9):661-670, 2001.

Rizzolatti G and Craighero L. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169-192, 2004.

Rizzolatti G, Sinigaglia C, and Anderson F. *Mirrors in the brain*. Oxford University Press, 2008.

Rueschemeyer S, Brass M, and Friederici AD. Comprehending prehending: neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, 19 (5):855-865, 2007.

Santi A, and Grodzinsky Y. fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage*, 51 (4):1285-1293, 2010.

Saygin AP, Wilson SM, Dronkers NF, and Bates E. Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia*, 42 (13):1788-1804, 2004.

Schubotz RI and von Cramon DY. Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Brain Research. Cognitive Brain Research*, 11 (1):97-112, 2001a.

Schubotz RI and von Cramon DY. Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *Neuroimage*, 15 (4):787-796, 2002b.

Schubotz RI and von Cramon DY. A blueprint for target motion: fMRI reveals perceived sequential complexity to modulate premotor cortex. *Neuroimage*, 16 (4):920-935, 2002c.

Shannon CE and Weaver W. *The mathematical theory of communication*. Chicago, IL: University of Illinois Press, 1949.

Shetreet E, Friedmann N, and Hadar U. An fMRI study of syntactic layers: sentential and lexical aspects of embedding. *Neuroimage*, 48 (4):707-716, 2009.

Shtyrov Y, Hauk O, and Pulvermüller F. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *European Journal of Neuroscience*, 19 (4):1083-1092, 2004.

Simmons WK, Reddish M, Bellgowan PSF, and Martin A. The Selectivity and Functional Connectivity of the Anterior Temporal Lobes. *Cerebral Cortex*, 20 (4):813-825, 2010.

Tettamanti M, Alkadhi H, Moro A, Perani D, Kollias S, and Weniger D. Neural correlates for the acquisition of natural language syntax. *Neuroimage*, 17:700-709, 2002.

Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, Fazio F, Rizzolatti G, Cappa SF, and Perani D. Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17 (2):273-281, 2005.

Tettamanti M, Manenti R, Rosa PAD, Falini A, Perani D, Cappa SF, and Moro A. Negation in the brain: modulating action representations. *Neuroimage*, 43 (2):358–367, 2008.

Tettamanti M, Rotondi I, Perani D, Scotti G, Fazio F, Cappa SF, and Moro A. Syntax without language: neurobiological evidence for cross-domain syntactic computations. *Cortex*, 45 (7):825-838, 2009.

Tettamanti M and Weniger D. Broca's area: a supramodal hierarchical processor? *Cortex*, 42 (4):491-494, 2006.

Toni I, de Lange FP, Noordzij ML, and Hagoort P. Language beyond action. *Journal of Physiology, Paris*, 102 (1-3):71-79, 2008.

Tourville JA, Reilly KJ, and Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, 39 (3):1429-1443, 2008.

Tranel D, Kemmerer D, Adolphs R, Damasio H, and Damasio AR. Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology*, 20 (3,4,5,6):409-432, 2003.

Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C, and Rizzolatti G. I know what you are doing. a neurophysiological study. *Neuron*, 31 (1):155-165, 2001.

Warrington EK and McCarthy RA. Categories of knowledge. Further fractionations and an attempted integration. *Brain*, 110 (5):1273-1296, 1987.

Willems R, Ozyurek A, and Hagoort P. When language meets action: the neural integration of gesture and speech. *Cerebral Cortex*, 17 (10):2322-2333, 2007.

Willems RM and Hagoort P. Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language*, 101 (3):278-289, 2007.

Wilson SM, Saygin AP, Sereno MI, and Iacoboni M. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7 (7):701-702, 2004.

Womble S, and Wermter S. A Mirror Neuron System for Syntax Acquisition. *Artificial Neural Networks - ICANN 2001*:1233-1238, 2001.

Zanuttini R. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford, UK: Oxford University Press, 1997.

de Zubicaray G, Postle N, McMahon K, Meredith M, and Ashton R. Mirror neurons, the representation of word meaning, and the foot of the third left frontal convolution. *Brain and Language*, 112 (1):77-84, 2010.