



Sentence comprehension: effectors and goals, self and others. An overview of experiments and implications for robotics

Anna M. Borghi^{1,2*}, Claudia Gianelli¹ and Claudia Scorolli¹

¹ Department of Psychology, University of Bologna, Bologna, Italy

² Institute of Sciences and Technologies of Cognition, National Research Council, Rome, Italy

Edited by:

Angelo Cangelosi,
University of Plymouth, UK

Reviewed by:

Gert Westermann,
Oxford Brookes University, UK
Giovanni Pezzulo, Consiglio Nazionale
delle Ricerche, Italy

*Correspondence:

Anna M. Borghi, Dipartimento di
Psicologia, Viale Berti Pichat 5, 40127
Bologna, Italy.
e-mail: annamaria.borghi@unibo.it

According to theories referring to embodied and grounded cognition (Barsalou, 2008), language comprehension encompasses an embodied simulation of actions. The neural underpinnings of this simulation could be found in wide neural circuits that involve canonical and mirror neurons (Rizzolatti et al., 1996). In keeping with this view, we review behavioral and kinematic studies conducted in our lab which help characterize the relationship existing between language and the motor system. Overall, our results reveal that the simulation evoked during sentence comprehension is fine-grained, primarily in its sensitivity to the different effectors we employ to perform actions. In addition, they suggest that linguistic comprehension also relies on the representation of actions in terms of goals and of the chains of motor acts necessary to accomplish them. Finally, they indicate that these goals are modulated by both the object features the sentence refers to as well as by social aspects such as the characteristics of the agents implied by sentences. We will discuss the implications of these studies for embodied robotics.

Keywords: sentence comprehension, embodied cognition, robotics, action, motor system, language, action goals, social cognition

INTRODUCTION

According to theories of embodied and grounded cognition (from here on EC theories), language is grounded in the sensorimotor system. In this sense, the same sensorimotor and emotional systems are supposed to be involved during perception, action and language comprehension. More specifically, language comprehension would involve an embodied simulation, whose neural underpinnings are to be found in wide neural circuits, crucially involving canonical and mirror neurons (Rizzolatti et al., 1996; Gallese, 2008). In cognitive neuroscience the notion of simulation has been defined differently (for a more detailed analysis of this, see Borghi and Cimatti, 2010; for a review, see Decety and Grezes, 2006). Here we define simulation, with Jeannerod (2007), as the offline recruitment (for instance, during language processing) of the same neural networks involved in perception and action. In addition, we qualify it, as did Gallese (2009), as an embodied and automatic mechanism, which allows us to understand others' behaviors. The automaticity of this process does not imply an intentional strategy to understand intentions and mental states. In keeping with these views, the underlying assumption of our work is that the activation of motor and sensorimotor cortices is not just a side-effect but effectively contributes to language comprehension. In this paper we review behavioral and kinematics studies conducted in our lab which help to characterize the relationship existing between language and the motor system (see also Scorolli et al., 2009). We will focus on studies utilising simple sentences composed for example by a verb and a noun. In the final part of the paper we discuss why we believe these studies have implications for embodied robotics. Further, we will claim that embodied robotics can contribute critically to psychology and neuroscience and can promote more detailed predictions on some critical issues.

THE EXPERIMENTS

SENTENCE COMPREHENSION, SIMULATION AND EFFECTORS

Several recent studies have provided evidence of the involvement of the premotor cortex in reading and hearing action words and action sentences (Aziz-Zadeh and Damasio, 2008). Tettamanti et al. (2005) conducted an fMRI study illustrating that a complex frontoparietal circuit is activated when presenting sentences describing actions performed with the mouth, the hand or the foot. Within this circuit a critical role seems to be assumed by Broca's area, but in a way that extends the traditional linguistic role of this area. In fact, Broca's area is found to be crucially involved in language processing, as well as in action observation. Pulvermüller et al. (2001) found topographical differences in the brain activity patterns generated by verbs referring to different effectors (mouth, legs, arms: e.g. lick, kick, pick); these differences emerged quite early, starting 250 ms after word onset. This very fast activation, its automaticity and its somatotopic organization render it unlikely that information is first transduced in an abstract format and later influences the motor system, as claimed by critiques of the embodied view. In particular, the early activation of the motor system strongly suggests that this activation is an integrant part of the comprehension process rather than only a by-product of it, or an effect of late motor imagery. Further studies utilising a variety of techniques (fMRI, MEG, etc.) support the hypothesis that action verb processing quickly produces a somatotopic activation of the motor and premotor cortices (e.g. Hauk et al., 2004; Pulvermüller et al., 2005). In line with these results, Buccino et al. (2005) designed a TMS study that showed an amplitude decrease of MEPs recorded from hand muscles when listening to hand-action related sentences, and from foot muscles when listening to foot related sentences. This confirms a somatotopic

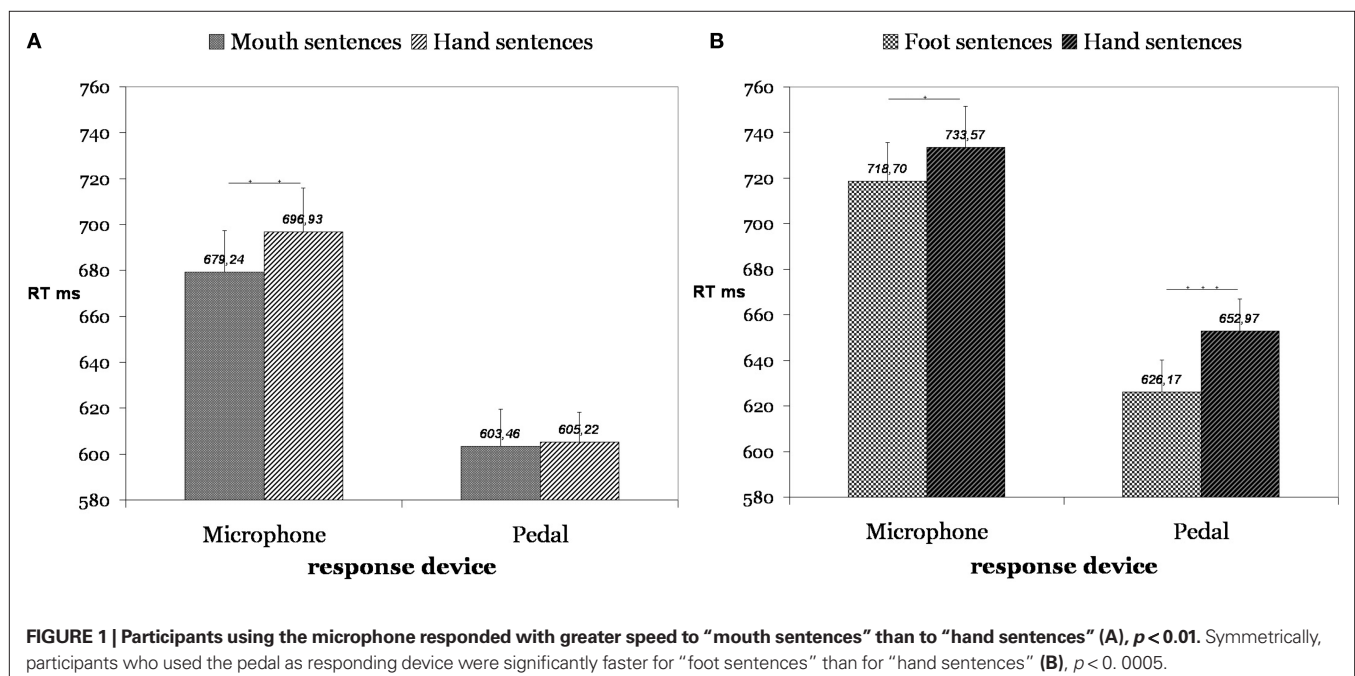
recruitment of motor areas. As reported in the meta-analysis performed by Jirak, Menz, Borghi and Binkofski (under review), the involvement of motor areas in language processing is consistent over tasks and subjects (for a more critical view, see Willems and Hagoort, 2007). In particular, word and sentence processing involves a variety of brain regions, including parietal, temporal, and frontal, but also cerebellar activity, and, even if the right hemisphere is also activated, there is a clear predominance of activations in the (language and motor areas of the) left hemisphere. In addition, the results of the meta-analysis highlight areas presumably containing mirror neurons in humans, more specifically Broca's region, which may be described as the human homolog of the monkey premotor cortex (Rizzolatti and Craighero, 2004).

We will now describe studies performed in our lab, as they extend the previous behavioral evidence. Here we have illustrated that during language comprehension we are sensitive to the distinction between hand and mouth sentences, and between foot and mouth sentences as well.

In the first study we performed two experiments in which 40 participants read simple sentences from a computer screen that were composed of a verb in the infinitive form followed by an object noun (Scorolli and Borghi, 2007). The sentences referred to either hand, mouth or foot actions. The hand sentences represented the baseline: thus, the same noun was presented after either a foot or hand verb (e.g. "to kick the ball", vs. "to throw the ball") or either after a mouth or hand verb (e.g. "to suck the sweet", vs. "to unwrap the sweet"). Overall, we had 24 object nouns, each preceded by two different verbs, for a total of 48 critical pairs. Presenting the same noun after the verb allowed us to be sure that no frequency effect took place. We did not control for the verb frequency, because the verb was presented before we started recording. However, in a pre-test we controlled for the association rate between the verb and the noun, as this might influence performance. Eighteen participants

were required to produce the first five nouns they associated with each verb; no difference in production means was present between "mouth sentences" and "hand sentences", $p = 0.65$, and between "foot sentences" and "hand sentences", $p = 1$. The timer started after the noun presentation, and participants were required to respond whether the verb–noun combination made sense or not. Yes responses were recorded either with the microphone or with a pedal. We found a facilitation effect in responses to "mouth sentences" and "foot sentences" compared with "hand sentences" when the effectors – mouth and foot – involved in the motor response and in the sentence were congruent (Figure 1). More specifically, participants responding with the microphone were faster with mouth than with hand sentences, $p < 0.01$ (Figure 1A), whereas the difference between foot and hand reached significance but was far less marked, $p < 0.05$ (Figure 1B). Participants using the pedal responded faster to foot than to hand sentences, $p < 0.0005$ (Figure 1B), whereas the difference between hand and mouth sentences was not significant, $p < 0.8$ (Figure 1A). These results suggest, in line with the literature, that the simulation activated during sentence comprehension is sensitive to the kind of effector implied by the sentence. In previous behavioral studies only foot and hand sentences were compared; our study extends previous results as we found a difference between mouth and hand sentences as well.

In a further study (Borghi and Scorolli, 2009) we found that the simulation is sensitive not only to the kind of effector (mouth vs. hand, foot vs. hand), but also to the specific effector (right vs. left hand) used to respond. We performed five experiments with the same sentence presentation modality and task used in Scorolli and Borghi (2007); 97 right-handed participants were asked to decide whether verb–noun combinations made sense or not. We analyzed both combinations which made sense (e.g. "to kick the ball") and combinations which did not make sense (e.g. "to melt the chair"). Here we will focus on Experiments 1, 2, and 3, as Experiment 4 was



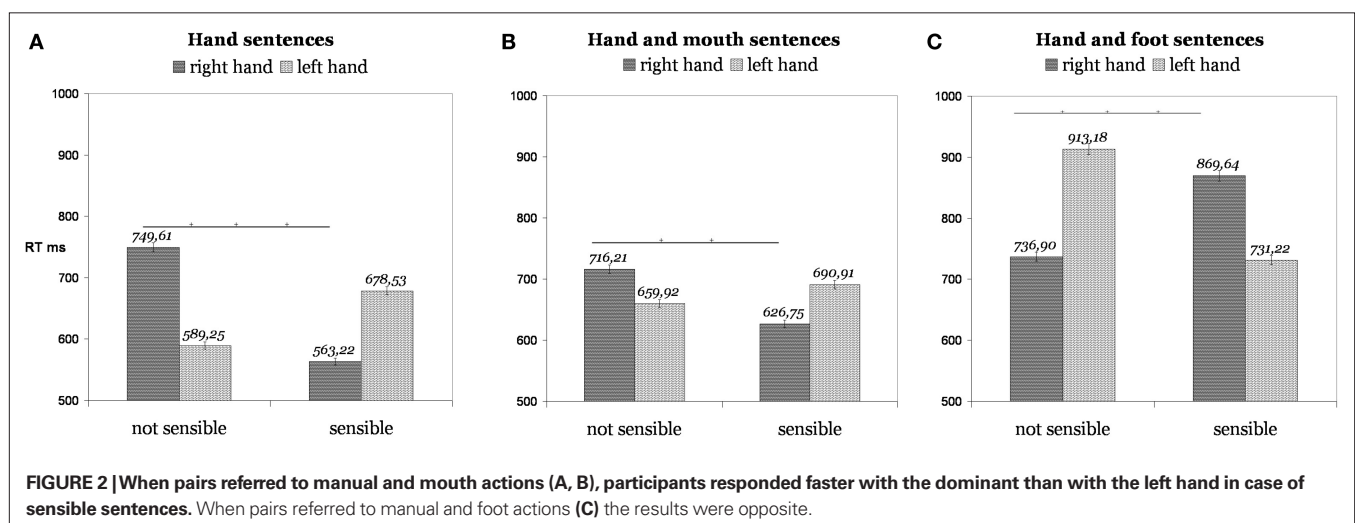
a control one. In Experiments 1a,b we used only manual sentences, in Experiment 2 hand and mouth sentences, in Experiment 3 hand and foot sentences. Responses to hand sentences (Experiment 1) were faster than responses to non-sense sentences with the right hand, but not with the left hand (**Figure 2A**), as it appeared in the subject analyses and on materials (we will report the p -values for both analyses in sequence): $p < 0.05$; $p < 0.0000001$. Importantly, such an advantage of the right over the left hand was not present when sensible sentences were not action ones: $p = 0.99$; $p = 0.75$. The same advantage of the right over the left hand with sensible sentences was present in Experiment 2 (**Figure 2B**), in which both hand and mouth sentences were presented, even if it reached significance only in the analysis on items, $p < 0.0000001$. This suggests that participants simulated performing the action with the dominant hand. Crucially the advantage of the right hand for sensible sentences was not present with foot sentences, with which, probably due to an inhibitory mechanism, the effect was exactly the opposite, as left hand responses were faster than right hand ones with sensible sentences, $p = 0.055$; $p < 0.0000001$ (**Figure 2C**). These results complement the previous findings as they suggest that the motor simulation formed is not only sensitive to different effectors (mouth, hand, foot), but also to the different action capability of the two hands, the left and the right one. The similarity between the responses with hand and mouth sentences can be due to the fact that different effectors can be involved in single actions, and the similarity of the performance obtained by hand and mouth sentences could be due to the fact that hands and mouth are represented cortically in contiguous areas. However, it may also suggest that not only proximal aspects, such as the kind of effector, modulate the motor responses, but also distal aspects, such as the action goal. Consider an action such as sucking a sweet: it probably also activates manual actions such as the action of grasping the sweet and bringing it to the mouth. In sum: it is possible that the similar modulation of the motor response is due to the common goal evoked by hand and mouth sentences (see also Gentilucci et al., 2008).

Overall, the results of these two studies indicate that language processing activates an action simulation that is sensitive to the effector involved. In addition, they suggest that understanding action sentences implies comprehension of the goals that the

actions entail. However, further studies are needed, to deepen the role played by action goals (for a recent study focusing on the importance of goals in action organization in monkeys, see Umiltà et al., 2008. The issue of goals will be discussed later). The results described so far report a facilitation effect in case of congruency between the effector implied by the verb/sentence and the effector used to respond. Even if the evidence we found supports the idea that the different effectors (mouth and foot) are activated during language processing, our behavioral results contrast with the results by Buccino et al. (2005), who found an interference effect between the effector involved in the sentence (hand, foot) and the effector involved in the motor response (hand, foot).

Certainly, in both cases there is clear evidence of a modulation of the motor system during sentence comprehension, thus this evidence is certainly in favor of an embodied cognition perspective. However, knowing more precisely the specific timing of this modulation (Boulenger et al., 2006), as well as the details of this modulation, would be crucial for solving a lot of issues. The first issue is that, even if the somatotopic activation of the motor system suggests that the motor system is involved during language comprehension, we do not yet fully understand if the activation of the motor system is necessary for comprehension or whether it is just a by-product of it (Mahon and Caramazza, 2008). A better understanding of the relationships between the comprehension process and motor system activation, both in terms of time-course and processes, would be crucial as it would allow researchers to formulate clearer predictions.

Many interpretations of the discrepancies between the results have been proposed. One possibility is that these discrepancies are due to timing between linguistic stimulus, motor instructions and motor response. It is possible that, when the motor system is activated both for preparing an action with a given effector and for processing action words referring to the same effector, an interference effect takes place due to the contemporary recruitment of the same resources. Later, a facilitation effect might occur (see Chersi et al., 2010). This explanation is in line with evidence on language and motor resonance that has shown that the compatibility effect between action and sentence (ACE, that is the facilitation effect) was present only when the motor



instruction was presented simultaneously to the beginning of the sentence rather than after sentence presentation (Borreggine and Kaschak, 2006; Zwaan and Taylor, 2006). In the study by Buccino et al. (2005) participants on presentation of a “go” signal had to respond to the second syllable of a verb preceding a noun; time was measured from this point. Instead in our experiments we didn’t use a “go” signal: we first presented a verb, then a noun, and started measuring after the appearance of the noun. Concerning the temporal relationship between language and the motor task, the linguistic stimulus appearance seems to affect not only the movement’s speed (reaction times) but also the overt motor behavior, as revealed by detailed analyses of movement kinematics (Boulenger et al. 2006; Dalla Volta et al., 2009). Boulenger et al. (2006) found that when contemporaneously processing language and executing motor tasks, action verbs hinder reaching movements. An interference effect occurred as early as 160–180 ms after word onset when participants started the response movement before word presentation (Experiment 1). On the contrary a priming effect became evident at about 550–580 ms after word onset when the word acted as go-signal for the response movement. Along this line, Dalla Volta et al. (2009) found that there is an early interference effect on the effective movement (kinematics measures) and a late facilitation, detectable through RTs analyses.

Another possibility is that the interference effect is not only aroused by timing but by the interaction between two factors: the temporal overlap and the so called “integrability”, that is the degree to which the perceptual input could be integrated into the simulation activated by language. For example, in studies where both sentences and perceptual stimuli were presented, when the perceptual stimuli were abstract and difficult to integrate, an interference effect occurred; otherwise a facilitation effect took place. The difficulty seems to rest on the shared contents between the percept and the simulation of the sentence, and on the temporal overlap (Kaschak et al., 2005; Borreggine and Kaschak, 2006). However, due to the difficulty of integration between perceptual and linguistic stimuli this explanation may be contradicted when accounting for the interference and facilitation effects occurring when using linguistic stimuli.

A further possibility is that these discrepancies are due to the varying paradigms and stimuli used. For example, in some cases tasks requiring superficial processing (e.g. lexical decision tasks) are employed, whereas in other cases tasks requiring deeper semantic processing are used (this position is supported by Sato et al., 2008). More specifically, even in the case of deep semantic processing, results may differ depending on the task at hand. For instance, whether the task requires evaluating the whole sentence (e.g. Scorolli and Borghi, 2007; Borghi and Scorolli, 2009, asked participants to evaluate the sensibility of the verb–noun combination) or the verb (e.g. Buccino et al. required participants to evaluate whether the action verb was abstract or concrete).

A final possibility which should be explored is that the effect emerges differently depending on the type of pronoun used to indicate the agent of the action. In this sense pronouns induce a specific perspective on action, which modulates the motor system. For example, we found that the simulation triggered by the pronouns “I” and “you” have a different effect on kinematics parameters of

action. In addition, it is possible that the third person pronoun (see Buccino et al., 2005) may partially activate a simulation, thus relying on more abstract processes.

Overall, further research is necessary to disentangle which mechanisms are underlying interference and facilitation effects. However, we believe that further experimental data are not sufficient. Namely, modeling could help to understand how the process might occur, and might be helpful to propose more detailed and clearer predictions for new experiments. Modeling could help us to understand whether interference and facilitation are two sides of the same coin, or whether they rely on different mechanisms (for an attempt to model interference and facilitation effects, Chersi et al., 2010).

SENTENCE COMPREHENSION, SIMULATION, GOALS AND SOCIAL ASPECTS

In the previous studies we have seen that during language processing we form a simulation sensitive not only to the specific effector, but also to the goal conveyed by the sentence.

Consider for example giving somebody an object: how and to what extent is the action of “giving” represented differently from the action of, say, holding the object? These two actions imply two different goals, and these different goals imply a different chain of motor acts. Namely, in order to hold an object we need to reach and then grasp it, whereas in order to give an object to someone else we need to reach and grasp it, as well as to give it to the other agent involved in the interaction. Thus, in order to pursue the goal it conveys, this “interactive” action implies a longer sequence of chained motor acts.

Goal-relatedness of action has recently received much attention, in particular since Fogassi et al. (2005) demonstrated studying the monkey parietal cortex that motor acts, such as “grasping”, are coded according to the specific action (e.g. “grasping for eating” vs. “grasping for placing”) in which these acts are embedded. Moreover, this coding is present both when the action is performed and when it is observed, that is a mirror mechanism is involved. The idea that actions have a chained organization has been extended to humans, in particular for what concerns action observation and understanding. Iacoboni et al. (2005) used fMRI to demonstrate the presence of a chained organization that differs depending on the intention of the agent. Other studies have been conducted, showing that impairment of chain organization might be linked to autism spectrum disorder (Cattaneo et al., 2007; Boria et al., 2009; Fabbri-Destro et al., 2009). However, no behavioral task has yet been conducted, demonstrating the importance of chained organization in the normal adult population. Additionally, to our knowledge the only study investigating the extent to which this chained organization is encoded in language was a kinematics study recently performed in our lab, by Gianelli and Borghi (Gianelli and Borghi: I grasp, You give: when language translates actions, submitted), in which we identified different components of verbs (for a similar approach, see Kemmerer, et al., 2008) and distinguished between action verbs (e.g. “to grasp”) and interaction verbs (e.g. “to give”). These two kinds of verbs, which differ both for their chained organization and for their goal (acting with an object vs. interacting with another agent), had a different impact on kinematics parameters. That is, participants’ response (e.g. reaching–grasping an object)

was modulated according to the typical kinematics involved by the actions described by action and interaction verbs. Namely, since interaction verbs describe the interaction with another person, the kinematics in response to interaction verbs is modulated according to an increased requirement for accuracy and precision. That is, the same act of reaching and grasping an object needs to be more accurate when performed in order to give the object to another person, hence performing an additional motor act. Specifically, the deceleration phase is longer. The same effect is found during processing of verbs referring to the same action. This suggests that the chained organization of actions according to more or less interactional goals is translated by language. This chained organization can be reactivated when the motor system is activated, thus similarly contributing to language processing.

The results of this study suggest that sentences referring to actions involving other people (e.g. giving something) are represented differently in comparison to sentences referring to actions involving a relationship between an agent and an object (e.g. holding something). However, this study did not allow us to disentangle whether the difference was due to the different chain of motor events involved in the two actions, or whether it was due to a difference in the social framework the two sentences referred to. To elaborate, do “grasp” and “give” differ at a motor level because of the chain they imply, and the different motor acts used, or do they differ because their “goal”, as defined not only by a sequence of motor acts but also by the social dimension in which the action is performed? Namely, in the case of “give” the presence of another person is implied, while in the case of grasp it is not. Hence, their goal and their value differ. Even if the action chain organization characterizes both the canonical and the mirror neuron system, it is possible that, depending on the social framework the sentence describes, there is a different involvement of these two systems. Now consider words referring to objects which differ in valence, to take an example, words such as “nice” or “ugly”. Literature on approach/avoidance movements has used a variety of behavioral studies to demonstrate that when we read positive words we are faster in producing a movement with our body; the opposite is true when we read negative words (e.g. Chen and Bargh, 1999; Niedenthal et al., 2005; van Dantzig et al., 2008; Freina et al., 2009).

We conducted three experiments to explore whether the triadic relation between objects, ourselves and other natural and artificial agents modulates the motor system activation during sentence comprehension. We used sentences that referred to nice/ugly objects and to different kinds of recipients (Lugli, Baroni, Gianelli, Borghini and Nicoletti (under review)). Participants were presented with sentences formed by a descriptive part (e.g. the object is attractive/

ugly) and by an action part (bring it towards you/give it to another person). Their task consisted of deciding whether the sentence made sense or not by moving the mouse towards or away from their body. In three experiments we manipulated the recipient of the action, which could be “another person”, “a table”, or “a friend”.

Results showed that the direction (away or towards the body) of the movement performed to respond was influenced by the direction of the motion implied by the sentence and the stimuli valence. Crucially, stimulus valence had a different impact depending on the relational context the sentence evoked (action involving another agent or just oneself). We found that, whereas participants tended to move the mouse towards their body when they had to judge actions referring to positive objects, with negative objects the movement varied depending on the action recipient. Namely, when dealing with negative objects participants tended to treat friends as themselves, being equally slow to attract negative objects and to offer them to friends. This was not the case for the recipient “table” and for indistinct “another person”. In **Table 1** we compare the effect on RTs of different recipients, “another person”, “a table” or “a friend” in the two conditions of giving positive or negative objects.

A further result is worth noting. The paradigm we used in this study allowed us to disentangle information provided by the verb and kinematics information related to the real movement participants were required to produce to respond. Namely, given the experimental design we used, in half of the cases there was a mismatch between the information conveyed by the verb (bring vs. give) and the movement to perform (towards or away from participant’s body). Our results showed that the role played by the verb, which defines the action goal, was more important than the role played by the kinematics of the movement. This is in line with the Theory of Event Coding (Hommel et al., 2001), according to which actions are represented in terms of distal aspects, an overall goal, rather than in terms of the proximal ones, and with neurophysiological studies showing that actions are represented in the brain primarily in terms of goals (e.g. Umiltà et al., 2008).

DISCUSSION

Overall, our results suggest that the simulation evoked during sentence comprehension is fine-grained, as it is sensitive both to proximal and to distal information (effectors and goals). Additionally, the results show that actions are represented in terms of goals and of the motor acts necessary to reach them. Finally, they indicate that these goals are modulated by the characteristics of both objects and agents implied by sentences: this is observed due to the difference between actions involving only the self in comparison to those involving others.

Table 1 | Mean response times (RTs, in milliseconds) in the “another person” – “table” – “friend” target/negative object condition and “another person” – “table” – “friend” target/positive object condition.

Experiment	“Another person” – “table” – “friend” target/negative Objects	“Another person” – “table” – “friend” target/positive objects	difference
Exp. 1 “Another”	1645	1634	11
Exp. 2 – “Table”	1834	1834	0
Exp. 3 – “Friend”	1662	1609	53

We believe that realizing a model of these experiments would be important for understanding the relationships between language and motor system. Namely, modeling could contribute to create a theory of their relationship, which is detailed and advances clear predictions. In this direction, models can help to integrate a variety of different empirical results, obtained with different paradigms and different techniques, within a common framework. However, it is important that models do not only replicate experimental studies, but rather provide general principles and generate predictions to be tested empirically.

One could ask which kinds of models can help to interpret experimental results as the described ones, and help to formulate novel predictions.

Simple feed-forward models are probably not sufficient, as they may not provide an adequate formalization for embodied theories. Namely, feed-forward models are endowed with an input and an output lawyer which strongly resembles the traditional sandwich of dis-embodied theories of cognition. A recurrent network would probably be more suitable to detect the reciprocal influence of perception and action.

On a general level, modeling should respect a variety of constraints (see Caligiore et al., 2009; Caligiore, Borghi, Parisi and Baldassarre (accepted)).

The first kind of constraints are the neurobiological ones. Namely, the model's neural system should be endowed with at least some crucial characteristics of the human neural system. In particular, the neural underpinnings of motor simulations formed during language comprehension are represented by wide neural circuits that – crucially – involve canonical and mirror neurons (Gallese et al., 1996; Rizzolatti and Craighero, 2004). Therefore, the model should be endowed with a simulated neural system which reproduces both canonical and mirror neurons. More specifically, the motor system of the model should be organized in such a way that chains of actions are implemented so that each sequence includes different motor acts, and is organized around goals. One exemplar model, that clearly describes this phenomenon, was presented by Chersi et al. (2005, 2006), who modeled the study by Fogassi et al. (2005) using a chain model. Additionally, the model has been extended to explain how intention understanding and mental simulation take place. We believe that this model could be extended to study whether a chained organization explains the differences between verbs and sentences, for example between action and interaction verbs. Other hierarchical action schemas have been suggested in the literature, for example by Botvinick et al. (2009), adopted a reinforcement learning hierarchical model. Botvinick (2008) reviews how hierarchical models of action are being more frequently referred to. This is probably due to the fact that the way in which how general and abstract action representation emerges from action components, and the role of prefrontal cortex in this process, is becoming an important issue for neuroscientific research. The second kind of constraints are “embodiment constraints”. Namely, it would be important to replicate the experiments using embodied models, i.e. models endowed not only with a brain which is similar to that of humans, but also with a body which is similar to ours. In sum: robots should be endowed with a sensorimotor system similar, at least in some respects, to a humans' sensorimotor system. Consider that assuming a strong embodied view would lead to the claim that, given the differences between robots and humans sensorimotor

system, embodied robotics models cannot contribute to provide an adequate account of human linguistic comprehension capabilities. We prefer to adopt a weaker embodied view. We propose that robotic models can strongly represent embodied theories of cognition. To elaborate, robotics could be a powerful instrument to explore the extent to which the similarity between the sensorimotor system of different organisms, artificial and natural, constrains the emergence of cognition, and the emergence of language comprehension abilities. In this respect, it might be critical to use robotic models of the sensorimotor system that differ at different degrees from the human one. This could contribute to determine the importance of embodiment theories: namely, it would allow researchers to better understand which aspects of humans' neural and sensorimotor system are critical and determine modifications in humans' behavior.

The third type of constraints, which are referred to as “behavioral constraints”, are directly linked with the capability of the model to reproduce and replicate the behaviors produced during the experiments. Having a model which respects the three constraints we outlined would facilitate formulating a synthetic and general theory of the relationships between language and the motor system. Namely, it could contribute to a synthesis effort thus identifying the crucial underpinnings of our behavior.

We believe a model that accounts for the constraints we have illustrated should be able to individuate general principles that combine important characteristics of the relationship between language and the motor system. In the behavioral studies we reported, the critical points which are worth modeling are the following:

The fact that

- during language comprehension the underlying motor and premotor cortices are activated;
- the motor system has a chained organization, and that this organization is encoded in language;
- actions, as well as words and sentences referring to actions, are encoded firstly in terms of distal aspects (overall goal), then of proximal ones (e.g. effectors);
- the different social framework in which the actions are inscribed can change the way in which the action is represented.

On this basis, a model should contribute in detail and explain:

- the time-course as well as the mechanisms underlying the interference and the facilitation effects occurring between effectors implied by action verbs/action sentences and the effectors used to provide a response;
- the mechanisms according to which the different number of motor acts involved in an action chain constrain the comprehension of different action verbs and action sentences;
- the mechanisms according to which, even if the length of a motor chain does not differ, action goals have influences on the comprehension of action verbs/action sentences and how this influences movement;
- the mechanisms according to which, even if the length of a motor chain does not differ the language referring to the presence of objects and/or of other organisms implies the activation of different neural mechanisms (e.g. canonical vs. mirror neurons) which differently affect behavior.

CONCLUSION

In conclusion: we believe that embodied robotics can greatly contribute to a better understanding of the mechanisms underlying the relationship between language and the motor system. We argue that roboticists and modelers should work alongside empirical scientists in order to improve abilities to construe models which do not only account for empirical results, but also formulate predictions that constrain and guide new experimental research.

REFERENCES

- Aziz-Zadeh, L., and Damasio, A. (2008). Embodied semantics for actions: findings from functional brain imaging. *J. Physiol. Paris* 102, 35–39.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Borghi, A. M., and Scorolli, C. (2009). Language comprehension and hand motion simulation. *Hum. Mov. Sci.* 28, 12–27.
- Borghi, A. M., and Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia*, 48, 763–773.
- Boria, S., Fabbri-Destro, M., Cattaneo, L., Sparaci, L., Sinigaglia, C., Santelli, E., Cossu, G., and Rizzolatti, G. (2009). Intention understanding in autism. *PLoS ONE* 4, e5596. doi: 10.1371/journal.pone.0005596.
- Borreggine, K. L., and Kaschak, M. (2006). The action-sentence compatibility effect: its all in the timing. *Cogn. Sci.* 30, 1097–1112.
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* 12, 201–208.
- Botvinick, M. M., Niv, Y., and Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280.
- Boulenger, V., Roy, A., Paulignan, Y., Deprez, V., Jeannerod, M., and Nazir, T. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *J. Cogn. Neurosci.* 18, 1607–1615.
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cogn. Brain Res.* 24, 355–363.
- Caligiore, D., Borghi, A. M., Parisi, D., and Baldassarre, G. (2009). “Affordances and compatibility effects: a neural-network computational model,” in *Connectionist Models of Behaviour and Cognition II: Proceedings of the 11th Neural Computation and Psychology Workshop*, eds J. Mayor, N. Ruh and K.

ACKNOWLEDGMENTS

This work was supported by the FP7 project ROSSI, Emergence of communication in RObots through Sensorimotor and Social Interaction, Grant agreement No: 216125. Thanks to people of the LARAL (<http://laral.istc.cnr.it>) and LOCEN groups and especially to people of the EMCO group (www.emco.unibo.it) for comments and discussion (and in some cases participation to) on these studies. A special thank to Kate Burke for English revision of the text, and to Fabian Chersi for comments.

- Plunkett (Singapore: World Scientific), 15–26.
- Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., and Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17825–17830.
- Chen, M., and Bargh, J. A. (1999). Consequences of automatic evaluations: to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull.* 25, 215–224.
- Chersi, F., Fogassi, L., Rozzi, S., Rizzolatti, G., and Ferrari, P. F. (2005). Neuronal chains for actions in the parietal lobe: a computational model. *Soc. Neurosci. Conf.*, 412.8.
- Chersi, F., Mukovskiy, A., Fogassi, L., Ferrari, P. F., and Erilagen, W. (2006). A model of intention understanding based on learned chains of motor acts in the parietal lobe. In *proceedings of the 15th Annual Computational Neuroscience Meeting* 2006, Edinburgh.
- Chersi, F., Ziemke, T., and Borghi, A. M. (2010). Sentence processing: Linking language to motor chains. *Front. Neurobot.* 4, 1–9.
- Dalla Volta, R., Gianelli, C., Campione, G. C., and Gentilucci, M. (2009). Action word understanding and overt motor. *Exp. Brain Res.* 196, 403–412.
- Decety, J., and Grèzes, J. (2006). The power of simulation: imagining one’s own and other’s behavior. *Brain Research* 1079, 4–14.
- Fabbri-Destro, M., Cattaneo, L., Boria, S., and Rizzolatti, G. (2009). Planning actions in autism. *Exp. Brain Res.* 192, 521–525.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667.
- Freina, L., Baroni, G., Borghi, A. M., and Nicoletti, R. (2009). Emotive concepts and motor responses: attraction or repulsion? *Mem. Cogn.* 37, 493–499.
- Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Soc. Neurosci.* 3, 317–333.
- Gallese, V. (2009). Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood. *Psychol. Research* 73, 486–498.
- Gallese, V., Craighero, L., Fadiga, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Gentilucci, M., Dalla Volta, R., and Gianelli, C. (2008). When the hands speak. *J. Physiol. Paris* 102, 21–30.
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.
- Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one’s own mirror neuron system. *PLoS Biol.* 3, e79. doi: 10.1371/journal.pbio.0030079.
- Jeannerod, M. (2007). *Motor Cognition: what actions tell the self*. Oxford: Oxford University Press.
- Kaschak, M. P., Madden, C. J., Theriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., and Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition* 94, B79–B89.
- Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., and Wiley C. (2008). Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang.* 107, 16–43.
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70.
- Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krath-Gruber, S., and Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Pers. Soc. Psychol. Rev.* 9, 184–211.
- Pulvermüller, F., Härle, M., and Hummel, F. (2001). Walking or talking? Behavioral and electrophysiological correlates of action verb processing. *Brain Lang.* 78, 143–168.
- Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *J. Cogn. Neurosci.* 17, 884–892.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141.
- Sato, M., Mengarelli, M., Riggio, L., Gallese, V., and Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language* 105, 83–90.
- Scorolli, C., and Borghi, A. M. (2007). Sentence comprehension and action: effector specific modulation of the motor system. *Brain Res.* 1130, 119–124.
- Scorolli, C., Borghi, A. M., and Glenberg, A. M. (2009). Language-induced motor activity in bimanual object lifting. *Exp. Brain Res.* 193, 43–53.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., and Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281.
- Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F. Jezzini, A., Gallese, V., and Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proc. Nat. Acad. Sci. U.S.A.* 105, 2209–2213.
- van Dantzig, S., Pecher, D., and Zwaan, R. A. (2008). Approach and avoidance as action effect. *Q. J. Exp. Psychol.* 61, 1298–1306.

Willems, R. M., and Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain Lang.* 101, 278–289.

Zwaan, R. A., and Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *J. Exp. Psychol. Gen.* 135, 1–11.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 December 2009; paper pending published: 12 February 2010; accepted:

27 April 2010; published online: 14 June 2010.

Citation: Borghi AM, Gianelli C and Scorolli C (2010) Sentence comprehension: effectors and goals, self and others. An overview of experiments and implications for robotics. *Front. Neurobot.* 4:3. doi: 10.3389/fnbot.2010.00003

Copyright © 2010 Borghi, Gianelli and Scorolli. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.