Anna M. Borghi. Action language comprehension, affordances and goals. In Yan Coello, Angela Bartolo (Eds). (in press early 2012). *Language and action in cognitive neuroscience. Contemporary topics in cognitive neuroscience series*. **Psychology Press.**

Anna M. Borghi*°

*Department of Psychology, University of Bologna; °Institute of Cognitive Sciences and Technologies, National Research Council, Rome

Action language comprehension, affordances and goals

Correspondence: Anna Borghi - Department of Psychology, Viale Berti Pichat, 5, 40127 Bologna – ITALY
Phone: +39-051-2091838, Fax: +39-051-243086, Email: annamaria.borghi@unibo.it
Web-site: http://laral.istc.cnr.it/borghi

Word count: abstract: 239 words; Text (including references): 8356

Abstract

So far studies on language comprehension adopting the embodied and grounded perspective have underlined the similarities between language and action. In line with the idea of neural re-use, this chapter shows that language reflects some characteristics of action organization, but also that it builds on these characteristics, and modifies them in sophisticated ways.

In the first part evidence is provided showing that the simulation evoked by language is rather detailed and it activates a specific kind of affordances, that is stable and canonical affordances rather than variable ones; in the case of tools this simulation activates function rather than manipulation. When possible, the simulation prepares for a specific action; when the situation is uncertain predictions are advanced in parallel in preparation of many possible situations.

The second part of the chapter discusses how language reflects some characteristics of action organization. Evidence is illustrated indicating that the goal-derived and chained structure of action is reflected in language. The simulation is influenced both by the overall action goal and by the kinematics of actions; the latter is relevant only if it leads to a different interpretation of the action as a whole. However, language modulates how these aspects influence the motor system. The verb presented in imperative form, and in the 2nd person, leads to the preparation of a specific action, as the reader is recruited as agent. This phenomenon is less marked with other past tenses and with other pronouns.

1. Introduction

What internal processes take place when we are having a conversation, listening to somebody, or when we are reading?

In this chapter I will overview recently performed studies, with the aim of showing that embodied simulation is involved in action language comprehension. I will characterize simulation arguing that it is a rather detailed form of re-enactment of previous perceptual and motor experiences, but that it also plays a predictive role as in action preparation.

First, I will show that during language comprehension objects affordances are recruited. I will also specify which kinds of affordances are reflected in language. I will argue that language preserves memories of our more frequent interactions with objects, keeping track of the more stable and of the functional aspects of objects.

Second, I will show that language reflects the dynamics of action organization (Claudia Gianelli, PhD dissertation, 2010). In particular, language reflects two important – and related - characteristics of our motor system, i.e. the fact that it is organized in terms of goals and that it has a hierarchical chained organization, as each action is composed of different motor acts.

I will refer to behavioural or kinematics studies in which we did not use purely linguistic stimuli. Rather, words or simple sentences (noun-verb or pronoun-noun combinations) were preceded or followed either by images or by an object with which to interact (e.g., a box to lift, a mouse to grasp).

In order to address these points, it is important to make a premise clarifying the notion of simulation and its role.

2.Simulation between reenactmentreenactment and prediction

According to embodied and grounded cognition views, when we comprehend a sentence we re-activate the sensorimotor experiences the sentence refers to – we activate a "simulation". The idea that language comprehension recruits the same perception, action and emotional systems involved in interaction with objects (Fischer & Zwaan, 2008) is based on the underlying "neural exploitation" hypothesis (Gallese & Lakoff, 2005; Gallese, 2008; Glenberg & Gallese, in press; see also Glenberg, 2008), a model of neural re-use (Anderson, 2010); basically the idea is that, given that evolution is conservative, language relies on previously formed systems, such as the action system. The neural basis for simulation can be represented by the canonical and mirror neuron system (Rizzolatti & Craighero, 2004). Even if the focus of this chapter will be on language comprehension, it has been recently shown that the mechanisms and the neural systems underlying language comprehension and production might be similar (D'Ausilio, Craighero & Fadiga, in press; Gentilucci, Dalla Volta, & Gianelli, 2008).

The notion of simulation (or "emulation", as in Grush, 2004) is not uncontroversial in literature, as different versions of it have been proposed (for a review, see Decety & Grezes, 2006). Here the term "simulation" is used to refer to a process that is embodied, unconscious and not deliberate , and that is aimed at action preparation (Gallese, 2009). Therefore the simulation does not depend on a deliberate reactivation of previously performed actions, and it does not consist in s a form of motor imagery which occurs a posteriori (for a different view, see Decety & Ingvar, 1990).

Indeed, the notion of simulation as intended here has two crucial aspects: it implies the re-enactment of past sensorimotor experiences (Barsalou, 1999), and it has a predictive character (Grush, 2004; Pezzulo & Castelfranchi, 2009) as it prepares for action. Simulation is a form of re-enactment; for example, when reading "open the door", the same perceptual, motor, and emotional systems are recruited, which are involved when we perform the action. However, the motor simulation evoked during sentence comprehension cannot be reduced to a form of re-enactment: it has an anticipatory and predictive character as well; for example it can help us to anticipate the characteristics of the mentioned objects. Thus, reading the sentence "Grasp the brush" might lead us to prepare a power grip (i.e. a grip suitable to hold an object using all digits in a palmar opposition) rather than a precision grip (i.e., a grip characterized by an opposition between the index finger and the thumb); similarly, hearing the sentence "Lift the pillow"

might induce us to prepare ourselves to lift a light and soft object, exerting a specific kind of pressure. This anticipation can prepare an explicit action or a covert action, as "simulating is not doing" (Jeannerod, 2007).

This however, is not the whole story. Let us assume that during language comprehension we both re-enact previous experiences and predict possible novel experiences in order to prepare ourselves for action, and let us assume that language exploits the same mechanisms developed for action. It is important to understand the extent to which the simulation evoked during language comprehension is similar and the extent to which it differs from the simulation triggered by observation. So far studies on embodied cognition have highlighted the similarity between the simulation evoked during observation of actions and of objects and the simulation elicited by language, for example in sentences referring to actions with objects. However, there might be important differences between comprehension processes (say, the comprehension of an action) mediated by observation or by language (Parisi, in preparation). For example, observation and real interaction with objects is always situated in a specific context, whereas during language comprehension we typically have to mentally construct a situation. I will address these differences in the remainder of the chapter.

3. Language and prediction: Stable and variable affordances

The notion of affordance, as initially proposed by Gibson (1979), referred to the fact that objects and entities in the environment afford and invite organisms to act. For example, observing a bottle would invite us to grasp it. This notion has been recently re-evaluated, strengthened by brain imaging evidence which shows that observing tools activates motor areas of the brain (for a review see Martin, 2007). Ellis and Tucker (2000) proposed a variation of this notion. They chose the name "micro-affordance" to underline both the similarities with, and the differences from, the original concept of affordances. Similarly to Gibsonian affordances, micro-affordances are elicited automatically, and facilitate simple interactions with objects. Differently from affordances, however, they pertain to specific action components, such as reaching and grasping; in addition, they are a consequence of object-based attention (Vainio, Ellis & Tucker, 2007). For example, a glass is represented by accessing the information that it can be reached and grasped with a specific kind of grip. A further difference is that Gibson was not interested in the brain processes underlying affordances; in contrast, micro-affordances can be conceived of as brain assemblies that represent objects, and that derive from previous visuomotor interactions with objects.

Recently, we proposed a distinction between stable and variable affordances (Borghi & Riggio, 2009). Stable affordances would emerge primarily from properties like size, which are rather constant across contexts; variable affordances would emerge primarily from properties such as orientation. A special case is represented by what we called "canonical" affordance. For example, the orientation of the handle of a cup can be considered a variable affordance, as it might change from one moment to another. Therefore it wouldn't make sense to store in memory information about it. However, even if we interact with cups in different orientations – for example when we wash them, move them, etc. -, when we use a cup it typically has a given orientation: it is upright, as we have to drink from it. Due to the higher frequency of this orientation when we use cups, it might be useful to store information on their canonical orientation.
At a neural level, preliminary evidence (Menz, Thill, Buccino, Borghi & Binkofski, in preparation) showed that stable affordances are represented more ventrally, whereas variable affordances are represented more dorsally. More specifically, even if there seems to be a partial overlap of the two kinds of affordances, the first would activate primarily the dorso-dorsal, the second the dorso-ventral pathway of our brain (Rizzolatti & Matelli, 2003).

The distinction between stable and variable affordances helps us to distinguish what happens when we interact with objects, and what happens when we listen to or read words.
There are a number of possibilities. One possibility is that the way in which words are comprehended is not grounded, i.e. it is not linked to objects' perceptual and motor characteristics; words would only be understood in terms of other words associated to them (Landauer & Dumais, 1997).

Another possibility is that the activation of the perception and action systems is the same in the two cases, when we interact with a glass and when we read or listen to the corresponding word. According to the indexical hypothesis (Glenberg & Robertson, 2000), words and sentences are linked to objects in the world, their referents, or to analogical representations as pictures or perceptual symbols (Barsalou, 1999). For example, the word "handle" refers to a handle, or to an analogical representation of it. Thus words that refer to objects would evoke perceptual and motor information relative to such objects. Borghi (2004) presented participants with sentences such as "The body extracts the book" followed by a word; their task was to decide whether the word referred to a part of the mentioned object or not. Parts could either afford an action or not: for example, the cover is a better affordance for the action of extracting a book than the page. Nouns referring to parts from which it was easy to derive affordances were processed more quickly and the combination was evaluated as the one which made more sense, even if there was no difference in semantic association between the affording and not affording word parts and the previous sentence. Results support the Indexical hypothesis; they show that the way we understand sentences is not explained by associative relations between words; rather, it is constrained by the affordances elicited by the objects words refer to. However, this may not be the whole picture.

A third possibility which I will try to advance here is that sentence understanding is tied to and constrained by object affordances (for a computational model of the relationship between affordances and compatibility effect, which considers language as well, see Caligiore et al., 2010), but that language recruits primarily some kinds of affordances. There could be important differences between producing a motor response to a sensorial stimulus, like a cup, and producing a motor response to a different sensorial stimulus, such as the word "cup".

Assuming that understanding words and sentences is influenced by object affordances, it becomes important to understand and specify which kinds of affordances are activated, for example in reading the word "cup". In order to prepare ourselves to interact successfully with cups, we should retrieve what we know about cups. Based on linguistic information, we should retrieve information related to the kind of grip that is typically evoked by cups, since a rather stable property of our interaction with cups is that, when we grasp the cup's handle, we use a power grip. In addition, we might activate canonical affordances. For example, when we use brushes they have a specific orientation, so it is possible that, when we understand a word like "brush", we reactivate brushes in that particular orientation.

Borghi and Riggio (2009) intended to investigate which kinds of affordances were activated during a recognition task (Stanfield & Zwaan, 2001). Participants were presented with sentences referring to observation or to action (e.g., look at vs. grasp the brush); after 400ms the sentence disappeared and was substituted by a picture of an object. They had to respond by pressing a different key whether the object represented the object mentioned in the sentence or not. The objects could elicit either a power or a precision grip (e.g., brush vs. pencil), and could have the canonical orientation (i.e., have the affording part in the lower part of the screen vs. on the higher part of the screen) or not. We found that, even if the verb frequency was the same, action verbs (which were presented in imperative forms) were processed faster than observation verbs, indicating that they induced the production of an action.
In addition, when the objects were presented in the canonical orientation, RTs were faster than when there was a mismatch between the canonical orientation and the current orientation of the object. This effect is a motor, not a visual one, as it reflected objects' orientation for use, not their visual orientation. Consider for example a brush: typically it lies horizontally when we grasp it, but it has an upright position when we use it to brush our hair.
The effect of the canonical orientation did not interact with the verb type, therefore we cannot determine whether the result was due to the formation of a motor prototype while reading the sentence or was simply due to the fact that, independently of the sentence, we represent objects in terms of their canonical orientation. Further studies are necessary to disentangle this issue.
The most interesting results concern false trials, i.e. the cases when the object in the sentence and in the picture did not correspond. In false trials, when the grip required by the object in the sentence and the one requested to grasp the object in the picture corresponded (for example, the word "brush" was followed by

the picture of another object requiring a power grip, such as a hammer), RTs were slower with action verbs than with observation ones, probably due to an inhibition of the motor system. This result suggests that when we read an action sentence we prepare ourselves to act with a given object, and that the object name activates the way we should grasp that object. In other words, we represent an object in terms of its size, and of the grip it evokes. This inhibitory effect is compatible with the TEC theory (Hommel et al., 2001), according to which if an event file is activated from two different sources (the linguistic task and the motor one), an inhibitory process takes place.

Overall, these results allow us to characterize the simulation formed and to understand which kinds of affordances we evoke while reading action sentences. They indicate that we form a rather detailed simulation. Similarly to the simulation built during object observation, it re-activates previous experiences and predicts future actions, preparing us to act. Specifically, it appears that we build a motor prototype, which includes stable affordances, such as those emerging from the objects' grip and from their canonical orientation, i.e. their orientation for use. Thus the way we understand sentences is linked to object affordances, rather than to associative relationships between words. However, the simulation formed while reading words differs from the simulation built while observing objects. Indeed, the first activates only stable and canonical affordances, i.e. it anticipates the object properties which have been frequently experienced and have a higher probability to be encountered. Language does not seem to capture the variable properties of objects, which have to be responded to during direct interaction with them.

4. Language and prediction: manipulation and functional affordances

Beyond the distinction between stable, variable and canonical affordances, there might be further distinctions between affordances. For example, different affordances might be related to object manipulation and to object use. We can interact with objects in multiple ways. When we put a knife on the table, we grasp it with a precision grip, while when we use it to cut something we hold it with a power grip. Tools represent a special case of objects, as they can evoke skilled actions related to their use as well as grasping actions consistent with their structure (see Borghi, 2005; see the distinction between volumetric and functional gestures in Bub, Masson and Cree, 2008). At a neural level, it has been proposed that affordances related to object use are represented more ventrally, with those related to manipulation represented dorsally (Young, 2006).
Jax and Buxbaum (2010) demonstrated the difference between "conflict objects", which evoke conflicting grips for manipulation and use, and not conflict objects, for which manipulation and use are not in contrast. They propose that the intention to act on an object triggers a competition between responses aimed at manipulating vs. using the object. The activation of manipulation responses is quicker, given that functional responses require activation of stored information; thus less time is necessary to activate grasp responses; in addition, when grasp is performed before use there is an interference effect, which is not present when use is performed prior to the grasp. In a recent study we addressed the distinction between manipulation and use without conflict objects, but using a picture of a torch, characterized by a structural separation between the graspable and the goal directed part (handle, beam). When participants had to process the object shape and the torch was switched on, we found a compatibility effect between the key to press (left, right) on the keyboard and the handle location. This suggests that they simulated holding the handle to use the torch. The compatibility effect was not present when participants had to decide the colour of the torch. This result complements the previous one as it suggests that affordances related to function emerge when objects are processed at a certain depth -for example, attention should be paid to objects shape, not to their color (Pellicano, Iani, Borghi, Rubichi, & Nicoletti, 2010; Tipper, Paul, & Hayes, 2006).
What are the consequences of making the distinction between these 2 kinds of affordances on language comprehension? When we read or listen to sentences, do we represent objects in terms of affordances relevant for manipulation or for use? Our hypothesis is that, when for example a sentence is presented, functional information (e.g., gestures to use an object for its intended purpose) is more activated than information related to object manipulation (e.g., gestures to pick up an object). This should happen because functional actions are more frequently associated with tools than other kinds of actions.
Two studies in preparation support these hypotheses.

Marino, Borghi, Buccino and Riggio (in preparation) used the same paradigm as Borghi and Riggio (2009). Sentences were followed by objects graspable either with a power or a precision grip, with the handle orientation compatible or incompatible with the key to press. RTs were faster when the sentence contained the verbs related to function rather than simple manipulation and observation, even if the frequency of the verbs did not differ.

In a further study (Costantini, Ambrosini, Scorolli & Borghi, in press) we presented pictures of a 3D room with everyday tools, such as bottles, which could be located in the peri- or extrapersonal space. Participants read function, action and observation verbs (e.g., "to drink", "to grasp", "to look at") and had to judge if the verb was compatible with the presented object. RTs were shorter for function and manipulation than for observation verbs, suggesting that objects are represented in terms of affordances. In addition, responses with manipulation and function verbs were faster when objects were in the peri- than in extrapersonal space. In contrast, with observation verbs there was no difference between peri- and extrapersonal space. This reveals that context is flexibly modulated by the activated information. Finally, the effect of the difference between peri- and extrapersonal space was more marked with function than with manipulation verbs. Thus objects located in the reachable space evoke gestures aimed at manipulating and, more crucially, at using them.

In sum: during language comprehension we not only re-activate past sensorimotor experiences, but implicitly formulate predictions aimed at actions. To prepare for more frequent actions we activate a motor prototype of the objects referred to by the sentences. This prototype includes stable and canonical affordances, as well as affordances related to object use. Our results lead to hypotheses on the difference between the predictions advanced when we understand language versus when we interact with objects.

When we are presented with sentences followed by pictures or objects, we first access stable and canonical affordances, then we verify whether our predictions fit with what we see. The order of appearance of stable and variable affordances might differ for a non linguistic task, in which online information (e.g., variable affordances) is activated first.

The same might be true for manipulation and function information. Given that with tools functional actions are more frequent, it makes sense to activate through language a motor prototype based on function. The story can differ in the case of natural objects, where there is no conflict between manipulation and function, simply because with natural objects, such as peaches and cats, function is not activated. Jax and Buxbaum (2010) have shown that in real interactions information related to structural characteristics is activated earlier than functional information. With language we predict that the competition between grasp and use (Jax & Buxbaum, 2010) works exactly the opposite way round. Language would activate primarily function, whereas in online task manipulation is first evoked, eventually followed by activation of functional knowledge (for new evidence showing that object function rather than object manipulation is reflected in semantic representations in the brain, see Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010).

5. Language and what is unpredictable: the case of weight

A study which clarifies how the predictive function of simulation might work was performed on object weight (Scorolli, Borghi & Glenberg, 2009). Weight is an interesting property, because it cannot be inferred from visual properties, such as object size and shape, but it can be fully determined only through kinematics and kinaesthetic properties, i.e. thanks to real interactions with objects. Participants were acoustically presented with sentences with a verb (lift) followed by a noun referring to a light vs. a heavy object (e.g., pill vs. chest). Then they had to physically lift with both hands two different boxes, which were perceptually identical but differed in weight, one being light and the other heavy. In this way we ruled out any possible influence of visual properties, such as shape and size. We analysed the kinematics of the lift delay, defined as the time immediately after the object is grasped; we focused on this phase as it is shaped by proprioceptive features that cannot be visually detected. Results showed that participants' time delay was larger when the weight suggested by the sentence and the actual weight of the box corresponded. The results were consistent with the MOSAIC model (Hamilton, Wolpert & Frith, 2004), according to which the force used to perform an action results from the integration of the force parameters from several modules

applicable in that context (e.g., modules for light vs. for heavy objects); the integration is based on the estimated probability that a module applies in the situation. In our case participants had to perform 2 tasks, let's call Task 1 the language comprehension task and Task 2 the box lifting one. The simulation formed during Task 1 occupied a given module (e.g., the module for lifting a light object), rendering it unavailable for the subsequent task. This demonstrates how simulation might work, and how it influences and constrains action. Simulation does not equate with the simple way in which language influences and controls action. Indeed, our results partially differ from those obtained on expectation. Results on monomanual lifting (Jenmalm, Schmitz, Forssberg, & Ehrsson, 2006) indicate that, when an unpredictable light weight follows a heavy weight, lift movements are faster: this finding was consistent with our results. However, when an unpredictable heavy weight is lifted after a light weight, the duration of the loading phase is longer than when a heavy weight is lifted after another heavy weight: this result was the opposite as the one we found.

At a closer look, our data differed in the first and second part of the study. For the first trials the pattern of our results matched the results found on action prediction; the story was different for the second part of the experiment. This suggests that, at first, participants used the information conveyed by language in order to prepare for action. At a certain point, they realized that the weight of objects mentioned in the sentence was not a good predictor of the weight of the box they had to lift. In different terms, participants realized that the information priors on which they relied, derived from their previous experience with language and objects, did not now apply, leading to uncertainty. From this moment on, language could not be used as a good predictor. Nonetheless, our data suggests that participants could not avoid simulating while comprehending language. In a case like this, when the information conveyed by language does not produce reliable predictions, simulating does not correspond to forming an expectancy on the weight of the object to lift. Rather, it seems that the simulation formed in these cases lasts longer, probably because it is not aimed at performing an immediately subsequent action. Therefore it occupies some resources (modules) as a prolonged action does, and, because it renders those resources (modules) temporarily unavailable, it influences performance on a subsequent task.

The study on lifting allows for better characterization of how a simulation is triggered by language. When we comprehend language, the default functioning of the system leads to prediction. However, a specific prediction is built only when it is effectively useful in preparing to interact with objects. In some cases, when the context is not sufficiently informative, it might be difficult to select a specific action to prepare, and multiple predictions might be advanced at the same time. When the context of action is not clearly determined, then the simulation can be better characterized as a form of re-enactment of previous experiences. This doesn't mean that no prediction is advanced, but that the system is not oriented towards the preparation of a specific action; rather, it is open to different possibilities. In all cases, simulation plays an anticipatory role, it is not a form of imagery occurring a posteriori.

6. Language and the dynamics of action: goals and kinematics

If neural circuits initially used for one purpose (e.g., motor control) are re-used for other purposes (Anderson, 2010), then the characteristics of action organization should be reflected in language, in line with the neural exploitation hypothesis. Important characteristics of how actions are represented in the brain are goal-derivedness and chained organization. Fogassi et al. (2005) found that in the monkey parietal cortices neurons respond differently to the overall action goal in comparison to single motor acts that compose each action. If language is grounded in the perceptual and motor systems, then it would be important to demonstrate that language reflects the chained organization of action.

At the same time, as proposed in the theory of event coding (Hommel et al., 2001), action is organized in distal rather than in proximal terms, i.e. the goals play a more important role than the way in which the action is performed. This organization of action in terms of goals has been highlighted also by recent studies on the mirror neuron systems (e.g., Umilta' et al. 2008). We investigated the role played by goals and by motor chains in action organization in some recent studies.

We referred to the literature on approach-avoidance effect; in particular, Chen and Bargh (1999) found that responses with positive words were facilitated when participants had to pull a lever, responses with negative words were faster when they had to push a lever away from their body. We were interested in verifying whether the simulation formed would be detailed enough to detect differences in hand posture, if the hand posture influenced the meaning of the action (Freina, Baroni, Borghi & Nicoletti, 2009). Participants responded whether words were positive or negative by pressing a button either near or far away from the body. When they had to hit the buttons with the hand open, participants were faster when pressing the button away from their body for positive words, and when moving towards their body for negative words, as if they simulated reaching for something positive and avoiding something negative. We found the opposite pattern when participants responded holding a tennis ball in their hand, as if they simulated keeping "good" things for themselves and throwing away "bad" ones. This study, in line with the literature, reveals that processing words with different valence evokes differently oriented reaching movements. Differently from other studies, however, it suggests that the simulation was sensitive to hand posture, and it reveals that the movements are perceived in terms of their effects. Our results cannot be accounted for in terms of a specific muscle activation explanation, according to which approach-avoidance effects are due to an association of flexion movements with positive stimuli and of extension movements with negative stimuli (Tops & de Jong, 2006). Rather, the same flexion or extension movement can be associated either to positive or negative stimuli depending on the overall action goal, which was rendered explicit by the kinematics of the hand. This highlights the dynamical and contextual dependent relationship between language and motor system, and it reveals that the simulation is sensitive both to the overall action goal and to fine-grained kinematics aspects, such as hand posture, provided that that they elicit a different interpretation of the whole movement (avoid- keep vs. reach-throw away) (Hommel et al., 2001).

7. Language and prediction: situatedness and linguistic framework

An important difference between the simulation triggered by language and during action observation concerns the way in which the linguistic framework of the sentence influences the simulation. I will refer first to the context evoked by the sentence, then to the linguistic framework of the sentence, such as the kind of verb, the verb tense and the perspective the sentence refers to.

A first difference is that real interaction with objects is always situated, whereas this is not the case for language. In particular, in many experiments words and sentences are presented without a supporting context. In the absence of a specific context, the simulation evoked by language prepares for more typical situations. To clarify this point, it is worth illustrating a study performed with a property generation task, in which participants were linguistically described 3 different scenarios (imagine to build, to use-act with vs. to observe an object) and a neutral scenario, in which the object nouns were simply mentioned (Borghi, 2004). The rated salience and the order of production of parts of complex artifacts such as cars and washing machines were predicted by their relevance for more typical actions. Across the situations the parts relevant for action/use were produced earlier than those produced in the building and vision situations. However, depending on the scenario, different parts were activated. For example, for "car" 'pedals' were dominant in the action/use perspective, 'transmission' in the build perspective, and 'windshield' in the vision perspective. Results revealed that, in a linguistic task, artefacts are firstly represented in terms of the most frequent actions relevant to them, however the activated information is flexibly modulated by the context (for discussion on conceptual stability and flexibility, see Borghi, 2005). Still, it remains true that there are characteristics, such as the current orientation of objects (variable affordances), that are not retained in language, due to their high variability, but that strongly influence action (for example, they determine the orientation of the hand while grasping an object).

Beyond the context evoked by the sentence, other factors influence the simulation triggered by language. Here, I will report an example to clarify. Many recent studies have shown that the simulation formed during language processing is sensitive to the effector involved in the action. There is evidence of both interference and facilitation when the effector (e.g., leg-foot, arm-hand, face-mouth, but also left and right hand) implied by the sentence and the one used for responding correspond. Both facilitation and

interference certainly demonstrate that language comprehension modulates the motor system; however, it is difficult to account for contrasting results. A close analysis of brain imaging and behavioural evidence (for a review, see Borghi, Gianelli & Scorolli, 2010) reveals that interference and facilitation might be modulated by a number of factors.

An important factor is the time in which sentence processing is measured. Very early recording typically leads to interference, late recording to facilitation. Facilitation results are typically taken by scholars that are not adopting an embodied perspective as evidence that the activation of the motor system might be simply a side-effect (Mahon & Caramazza, 2008; Toni, de Lange, Noordzij, & Hagoort, 2008). We have recently shown, using a computational model based on motor chains, that the two phenomena might be the two faces of the same coin (Chersi, Thill, Ziemke & Borghi, 2010): with early recording the motor chain of reaching and pressing the key to respond overlaps with the motor chain activated by the verb (e.g., reaching and grasping an object); the overlap between the two motor chains is not present with late recording.

More crucial to the point I am making here, the simulation might be strongly influenced by factors related to the way in which the linguistic framework is built, such as the kind of verb and the adopted perspective, as evoked by the pronoun (Gianelli, 2010).

Evidence has shown that the motor prototype, which is dynamic as it is continuously updated due to new interactions with novel category members, is activated particularly when the object noun is preceded by an action verb, not when it is preceded by a different kind of verb. In addition, the verb tense plays a role in influencing the simulation. When an imperative form is used, then the reader or the recipient is typically called into play as an agent, and a specific action preparation process might start. This is not the case when an imperfect verb form is used ("She sewed the skirt", Buccino et al., 2005); the cases in which an infinitive tense is used can be considered as an intermediate situation (e.g., "to kick – ball").

Finally, a crucial role in constraining the simulation is played by perspective. Anticipation might be stronger, and lead to a facilitation rather than to an interference effect, if the reader/listener is recruited directly, leading him/her to adopt the agent perspective (through the use of the 2nd person pronoun). Otherwise, the sentence is interpreted as describing a situation, or telling a story, in which no specific agent is called into play. When this happens, a simulation is run, but in this case the re-enactment aspect is more relevant than the predictive one: we internally reproduce the situation, without preparing ourselves to act. From analysing the literature (see Gianelli, 2010, for a thorough analysis of this) we can see that, in studies in which an interference effect between the effectors involved in the sentence and in the motor response is obtained, the sentences are typically presented in the first or third perspective (e.g., Buccino et al., 2005); otherwise typically a facilitation is found (Gianelli & Borghi, under review).

8. Language, motor chains and perspective

Evidence revised so far shows that language builds on a previous system, the action system, and shares with it an underlying organization. For example, it is structured in terms of goals, and it has a chained structure. However, these characteristics are modulated by constraints that pertain the linguistic context. A study performed in our lab (Gianelli & Borghi, under review) focuses on how the dynamic interplay of long term goals and single motor acts that characterize action organization is reflected in action, and how this interplay is constrained by the way in which perspective is encoded in language. Consider two different verbs characterized by a different componential structure, such as "to grasp" and "to offer" (for a similar approach, see Kemmerer, Castillo, Talavage, Patterson, & Wiley, 2008): the first verb implies a dual relationship between an agent and an object/entity, while the second refers to a triadic relation between an agent, an object and another organism/patient. The two verbs are characterized by different motor chains: grasping an object would imply reaching for it and grasping it, whereas offering it to somebody else would imply a longer motor chains characterized at least by a further motor act. What happens when we read verbs of these two different kinds, anticipated by a pronoun?

In our kinematics study participants had to read sentences (a pronoun, I, You, or He followed by an action or an interaction verb) while reaching and grasping a mouse; once the sentence was completed, if the pronoun-verb combination was correct, they had to move the mouse away from their body. This paradigm

was aimed at studying how the perspective induced by different pronouns and the chained structure of different kinds of verbs had an effect on overt movement. For this reason we recorded and then analysed the kinematics of their reaching and grasping movements. As regards the role of the pronouns, two alternative predictions can be advanced. The first is that, while reading sentences, participants could imagine being involved in a conversation. In conversations the agent is typically recruited with the YOU pronoun. If participants imagined being involved in a conversation, then the second person pronoun should lead them to feel recruited as agents. Therefore the YOU pronoun should lead them to simulate the described actions more than the two other pronouns, particularly of the third person (HE), who would be perceived as an external participant.

Alternatively, participants could activate the content of the sentences, as if they were reading a book. In this case, there should not be an advantage of the second person; instead, participants could either shift in perspective, being equally sensitive to all the three perspectives, or they could simulate more when the first person pronoun was presented.

Our results reveal that the perspective expressed by the sentence influences kinematics parameters, emerging from the very early stages of movements. More specifically, results directly highlight the role played by the YOU pronoun in taking the agent's perspective. When the agent perspective was taken (with the YOU pronoun), motor programming was accomplished earlier with action verbs than with interaction verbs. The difference between the two verbs might be due to the different length of the motor chains they imply. Indeed, both action and interaction verbs imply the act of grasping an object. However, compared to action verbs, interaction verbs imply a further motor act, consisting in giving the object to somebody else. Differently from the YOU pronoun, the first and the third person pronouns did not differently activate the two kinds of verbs. This finding marks a difference between the simulation evoked by observation and by language. Many studies have shown an advantage of the first perspective when an action is observed. A specificity of the linguistic situation is that, because of its intrinsically social character and because the framework of a conversation is activated, the agent role is taken with the YOU pronoun.

Overall, these results indicate that the chained organization characterizing the motor system is encoded in language. In line with the re-use hypothesis, the linguistic system re-uses basic aspects of the action system. However, these aspects are filtered and modulated in a completely novel way, due to the specific characteristics of the linguistic medium. For example, the way in which pronouns lead to the adoption of a given perspective modulates the motor system leading to differences in the activation of the chained structure of different verbs.

Conclusion

The framework underlying this chapter is the embodied and grounded cognition view (e.g., Barsalou, 2008). So far studies on language comprehension adopting this perspective have underlined the similarities between language and action organization and structure. In line with the idea of neural re-use, in the present paper it has been shown that language reflects some characteristics of action organization. At the same time, however, language builds on these characteristics, and modifies them in multiple ways.

The first part of the chapter discusses how the simulation evoked by language reflects object affordances and it shows that language recruits preferentially some kinds of affordances. The data suggest that the simulation formed is detailed as it contains information on objects' typical size, orientation, and weight. During language comprehension a motor prototype is built which includes stable and canonical affordances rather than variable ones. This motor prototype refers to functional aspects, in particular when the noun refers to a tool (e.g., we activate how to grasp a hammer to use it, rather than simply to place it somewhere else). A critical case is represented by the situations in which there is a mismatch between the affordances of the simulated object and those of the presented object. In these situations a cost is paid, as it appears in the experiments on recognition and on box lifting. These cases allow us to capture the predictive character of the simulation. When the situation is really uncertain (for example, when we realize that there is the same probability that either a light or a heavy box might appear) we cannot prepare a specific action. This does not mean that no prediction occurs, but that we need to be more flexible; many predictions are activated in parallel, in preparation of many possible situations.

The second part of the chapter discusses how language reflects some characteristics of action organization and how it modifies and constraints them. To what extent is the dynamics of action organization reflected in language?

Our results on emotional stimuli suggest that the simulation is influenced both by the overall action goal and by the action kinematics, but the latter is relevant only if it leads to a different interpretation of the action as a whole. In addition, our results on action and interaction verbs indicate that verbs characterized by different motor chains have a differential influence on movement. Thus the goal-derived structure and the chained structure of action are reflected in language. However, language filters and modulates how these aspects influence the motor system. For example, when we hear or read a verb in imperative form, and in the 2nd person, we tend to predict and to prepare for an immediate, specific action. This phenomenon is less marked when the verb is in infinitive form or in the past tense, and when we are not directly recruited as agents through the YOU pronoun. Overall, language relies on the action system, reflects some of its characteristics, but filters and modifies them in sophisticated ways.

Among many others, one important aspect of relationship between language and action hasn't been treated here. Words are important not only as they refer to objects and situations, but also because they are tools aimed at acting in the world and at modifying it (Borghi & Cimatti, 2010; under review). This has interesting implications as it might suggest that words, similarly to tools, modify the perception of our body, as well as of the world surrounding us.

References

Anderson, M.L. (2010). Neural re-use as a fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33,* 245-266.

Barsalou, L. W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences,* 22**,** 577-609.

Barsalou, L.W. (2008). Grounded Cognition. *Annual Review of Psychology*, 59, 617-645.

Borghi, A.M**.** (2004). Objects concepts and action: Extracting affordances from objects' parts.*Acta Psychologica, 115, 1,* 69-96.

Borghi, A.M. (2005). Object concepts and action. In D. Pecher & R.A. Zwaan (Eds). *Grounding Cognition: The role of perception and action in memory, language, and thinking.* Cambridge: Cambridge University Press.

Borghi, A.M**.**, & Cimatti, F.(2010). Embodied cognition and beyond: Acting and sensing the body. *NeuropsychologiA, 48*, 763-773.

Borghi, A.M., & Cimatti, F. (under review). Words are not only words: the social acquisition of abstract words.

Borghi, A.M**.,** Gianelli, C., & Scorolli, C. (2010). Sentence comprehension: effectors and goals, self and others. An overview of experiments and implications for robotics. *Frontiers in neurorobotics.*

Borghi, A.M**.,** & Riggio, L. (2009). Sentence comprehension and simulation of objects temporary, canonical and stable affordances. *Brain Research, 1253,* 117-128.

Bub, D.N., Masson, M.E., & Cree, G.S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. Cognition. 106:27-58.

Bub, D., & Masson, M. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition, 116,* 394-408.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. Cognitive Brain Research, 24, 355-363.

Caligiore, D., Borghi, A.M., Parisi, D., & Baldassarre, G. (2010). TRoPICALS: An Embodied Neural-Network Model of Experiments on Compatibility Effects. Psychological Review, 117, 1188-228.

Chen, M., & Bargh, J.A. (1999). Consequences of Automatic Evaluation: Immediate Behavioral Predispositions to Approach or Avoid the Stimulus. *Personality and Social Psychology Bulletin, 25,* 215-144.

Chersi, F., Thill, S., Ziemke, T., & Borghi, A.M. (2010) Sentence processing: linking language to motor chains. *Frontiers in Neurorobotics* 4:4. doi: 10.3389/fnbot.2010.00004.

Costantini, M., Ambrosini, E., Scorolli, C. & Borghi, A.M. (in press). When objects are close to you: object manipulation and use in the peripersonal space. *Psychological Bulletin & Review*.

Decety, J., & Grèzes, J. (2006). The power of simulation : Imagining one's own and other's behavior. *Brain Research, 1079,* 4-14.

Decety, J., & Ingvar, D.H., (1990). Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. *Acta Psychologica, 73*, 13–24.

D'Ausilio, L., Craighero, L., & Fadiga, L. (in press). The contribution of the frontal lobe to the perception of speech. *Journal of Neurolinguistics*.

Ellis, R., & Tucker, M. (2000). Micro-affordance: the potentiation of components of action by seen objects. *British Journal of Psychology*, *9*, 451-471.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language – A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology,* 61(6), 825 – 850.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, 308, 662–667.

Freina, L., Baroni, G., Borghi, A.M., & Nicoletti, R.(2009).Emotive Concept-Nouns and Motor Responses: Attraction or Repulsion? *Memory and Cognition, 37*, 493-499.

Gallese V**.**, & Lakoff G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Reason and Language. *Cognitive Neuropsychology, 22,* 455-479.

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social Neuroscience, 3*, 317-333.

Gallese, V. (2009). Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood. *Psychological Research, 73*, 486-98.

Gentilucci, M., Dalla Volta, R., & Gianelli, C. (2008). When the hands speak. *Journal of Physiology-Paris102*, 21-30.

Gianelli, C. (2010). *The language of action. How language translates the dynamics of our actions*. PhD dissertation.

Gianelli, C., & Borghi, A.M. (under review). I grasp, You give: when language translates actions.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Glenberg, A.M. (2008). Toward the integration of bodily states, language, and action. In G.R. Semin & E.R. Smith (Eds.). *Embodied grounding.* (pp. 43-70). Cambridge: Cambridge University Press.

Glenberg, A.M., & Gallese, V. (in press). Action-based Language: A theory of language acquisition, comprehension, and production. *Cortex*.

Glenberg, A.M., & Robertson, D.A. (2000). Symbol grounding and meaning: a comparison of high dimensional and embodied theories of meaning. *Journal of Memory and Language, 43*, 379-401.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences, 27*, 377–442.

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your Own Action Influences How You Perceive Another Person's Action. *Current Biology*, *14*, 493-498.

Hommel,. B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioural and Brain Sciences, 24*, 849-878.

Jax, S.A., & Buxbaum, L.J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition, 115*, 350-355.

Jeannerod, M. (2007). *Motor Cognition: What actions tell the self*. Oxford: Oxford University Press.

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 & Language*, 107, 16-43.

Jenmalm, P., Schmitz, C., Forssberg, H., & Ehrsson, H.H. (2006) Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. *Journal of Neuroscience, 26*, 9015–9021.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211- 240.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology (Paris),* 102, 59–70.

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

Menz, M., Thill, S., Jirak, D., Buccino, G., Borghi, A.M., Ziemke, T., & Binkofski, F. (in preparation). Neural pathways of stable and variable affordances.

Pezzulo, G. & Castelfranchi, C. (2009). Thinking as the Control of Imagination: a Conceptual Framework for Goal-Directed Systems. *Psychological Research, 73*, 559-577.

Pellicano, A., Iani, C., Borghi, A.M., Rubichi & Nicoletti, R. (2010). Simon-like and functional affordance effects with tools: The effects of object perceptual discrimination and object action state. *Quarterly Journal of Experimental Psychology, 63,* 2190 — 2201.

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

Rizzolatti, G., & Matelli, M., (2003). Two different streams form the dorsal visual system: anatomy and functions. *Experimental Brain Research, 153*, 146-157.

Rueschemeyer, S.A., van Rooij, D., Lindemann, O., Willems, R.M., & Bekkering, H. (2010). The function of words: distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience, 22*, 1844-51.

Scorolli, C., Borghi, A.M**.**, & Glenberg, A.M. (2009). Language-induced motor activity in bimanual object lifting. *Experimental Brain Research, 193,* 43-53.

Stanfield, R.A., & Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*, 153-156.

Tipper, S.P., Paul, M., & Hayes, A. (2006) Vision-for-action: The effects of object property discrimination and action state on affordance compatibility effects. *Psychonomic Bulletin & Review*, *13*, 493-498.

Toni, I., de Lange, F. P., Noordzij, M. L., & Hagoort, P. (2008). Language beyond action. *Journal of Physiology (Paris), 102,* 71–79.

Tops, M., & de Jong, R. (2006). Posing for success: Clenching a fist facilitates approach. *Psychonomic Bulletin & Review, 13*, 229-234.

Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F. Jezzini, A., Gallese, V., & Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Science, 105*, 2209–2213.

Young, G. (2006). Are different affordances subserved by different neural pathways? *Brain & Cognition, 62*,134-142.

Vainio, L., Ellis, R., & Tucker, M. (2007). The role of visuospatial attention in action priming. *The Quarterly Journal of Experimental Psychology*, *60*, 241-261.