**Published on Brain Research, 2009, vol. 1253, pp. 117-128**

# Sentence comprehension and simulation of object temporary, canonical and stable affordances

Anna M. Borghi° & Lucia Riggio*

°Department of Psychology, University of Bologna, Italy

*Department of Neuroscience, University of Parma, Italy

Pages headline: Sentence comprehension and affordances

Corresponding author:

Anna M. Borghi - Department of Psychology, University of Bologna

Viale Berti Pichat, 5, 40127 Bologna

phone: +39-051-3091838, fax: +39-051-243086 –

e-mail: annamaria.borghi@unibo.it

web-site: http://gral.istc.cnr.it/borghi

Acknowledgments

Thanks to Giovanni Buccino, Art Glenberg and Claudia Scorolli for discussion on this work and to Fabio Benatti for his help in preparing the stimuli. This work was supported by MIUR (Ministero Italiano dell’Istruzione, dell’Università e della Ricerca), Progetto PRIN 2006 and by the FP7 project ROSSI, Emergence of communication in RObots through Sensorimotor and Social Interaction, Grant agreement no: 216125.

ABSTRACT

Two experiments investigate the effects of language comprehension on affordances. Participants read a sentence composed by either an observation or an action verb (Look at / Grasp) followed by an object name. They had to decide whether the visual object following the sentence was the same as the one mentioned in the sentence. Objects graspable with either a precision or a power grip were presented in an orientation affording action (canonical) or not. Action sentences were faster than observation sentences, and power grip objects were faster than precision grip objects. Moreover, faster RTs were obtained when orientation afforded action. Results indicate that the simulation activated during language comprehension leads to the formation of a “motor prototype” of the object. This motor prototype encodes information on temporary/canonical and stable affordances (e.g., orientation, size), which can be possibly referred to different cognitive and neural systems (dorsal, ventral systems).



Keywords: language grounding – affordances – mirror neurons – canonical neurons - embodiment

## 1. INTRODUCTION

A recent body of work has revealed that words are not linked in an arbitrary way to their referents but are grounded in perception, action and in sensorimotor processes. According to the "embodied" theory of language comprehension, understanding a sentence regarding an object would entail a mental simulation of the situation the sentence describes. This implies that the same neural areas are recruited as those involved during perception and interaction with the object (Barsalou, 1999; Barsalou, Simmons, Barbey, & Wilson, 2003; Gallese & Lakoff, 2005; Glenberg, 1997; Glenberg & Robertson, 2000; Pecher & Zwaan, 2005; Pulvermüller, 2005; Zwaan, 2004). Much recent evidence obtained with response time studies (Borghi, Glenberg & Kaschak, 2004; Borreggine, & Kaschak, 2006; Boulenger, Roy, Paulignan, Deprez, Jeannerod, & Nazir, 2006; Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005; Scorolli & Borghi, 2007), with kinematic measures (Gentilucci & Gangitano, 1998; Glover & Dixon, 2002; Glover, Rosenbaum, Graham, & Dixon, 2004; Nazir, Boulenger, Roy, Silber, Jeannerod, & Paulignan, 2008), with eye tracking studies (Huettig & Altmann, 2005; Spivey & Geng, 2001; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), and with brain imaging studies (Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Kellenbach, Brett, & Patterson, 2003; Kemmerer, Castillo, Talavage & , Patterson, Wiley, in press; Pulvermüller, 2003), suggests that words evoke perceptual and motor information regarding their referents. In particular, words, like visual stimuli, evoke object affordances (Gibson, 1979). Affordances are what the environment offers acting organisms. They pertain to both perception and action. In addition, they are neither properties of the object / environment nor of the acting organisms. Instead, they are intrinsically relational properties. For example, a door handle affords opening for us, but not for a baby. In addition, a handle does not afford acting for a tree, which is not an acting organism. Thus, affordances are unique to a particular organism, to his/her body structure and bodily characteristics.

Various studies have shown that visual objects automatically evoke affordances (e.g., Creem-Regehs, Dilda, Viccrilli, Federed, and Lee, 2007). For example, Martin, Wiggs, Ungerleider and

Haxby (1996) have demonstrated with PET that naming tools selectively activates a left premotor area which is typically involved while imagining grasping an object with the hand, as well as an area in the middle temporal gyrus that is involved in producing action words. On the behavioural side, much of the evidence showing that seeing objects activates affordances is obtained with compatibility paradigms (see for example Tipper, Paul, & Hayes, 2007). However, these studies have not clarified the hypothesis that there might be different kinds of affordances. Consider our interaction with an object. In order to grasp effectively an object we have to know how and where to grasp it. In the first case, we can refer to stable features of the object, such as its size and shape, which can be incorporated into an object representation, stored in memory; but in the second case, we have to refer to more temporary aspects that depend on the way in which the object is shown.

Tucker and Ellis (1998) asked participants to decide whether objects were upright or reversed. They found a compatibility effect between the handle orientation of objects, which was not relevant to the task, and the position of the response key (left, right). Their results show that seeing an object with a left- or a right- oriented handle activated object affordances, leading to the facilitation of responses with the ipsilateral hand. The result has been replicated by Phillips and Ward (2002) with a priming paradigm. They presented participants with a visual object prime such as a frying pan with a left-or right-oriented handle, oriented towards or away from the participants, or in a neutral position. The prime was followed, after a varying SOA, by an imperative target requiring a response with the left or right hand or a foot-press. They found a compatibility effect between handle orientation and the key to press. This effect increased with SOA and slightly increased when the handle pointed towards rather than away from the observer. These results demonstrate that visual affordances (e.g., the object handles) potentiate lateralized responses corresponding with a given orientation of the affordances.

In our framework, orientation can be considered an example of a temporary affordance. Namely, orientation varies depending on the object's visual presentation. It does not represent a permanent characteristic of the object. For example, we might know that frying pans have a handle that we must grasp in order to use them, but information concerning the current orientation of the handle has to be processed online. However, there might be some kinds of orientation that are associated with the canonical interaction with and/or use of the object (Palmer, Rosch & Case, 1981; Riddoch, Humphreys, Hickman, Clift, Daly, Colin, 2006). Thus, we will define orientation in general as a temporary affordance, but the typical orientation with which we interact with objects – for example, the typical orientation with which we read a book - as a canonical affordance. Thus, we claim that there might be different kinds of affordances: stable affordances, such as shape and size, and temporary affordances, such as orientation. Within temporary affordances, there might be canonical and non-canonical affordances: that is a given orientation might be more typical than others.

Similar compatibility effects have been found with stable affordances such as object size. Ellis and Tucker (2000) found a compatibility effect between object size and the kind of grip used to respond whether or not the objects were artefacts or natural objects. Participants responded faster in the case of objects graspable with a power grip by mimicking a power grip with a device, and to objects graspable with a precision grip by mimicking a precision grip. Borghi, Bonfiglioli, Lugli, Ricciardelli, Rubichi and Nicoletti (2007) found that priming a hand shape (power, precision) facilitated responses in the case of objects graspable with the same kind of grip, provided that participants first were trained to associate their own movements with the postures of the priming hand. These results suggest that seeing objects of different sizes activated information on how to grasp them (precision or power grip), even if this information was not relevant to the task at hand. The compatibility effects are due to information stored in memory rather than to online processing of visual information; the fact that the result obtained by Tucker and Ellis has been replicated in an experiment using words serves as further confirmation (Tucker & Ellis, 2004). Ellis and Tucker refer to the potentiated elements of an action as “microaffordances”. Microaffordances are brain assemblies that represent objects; they are the product of the conjoining, in the brain, of visual responses and action- related responses that have developed throughout individual and species history, i.e. through both ontogenesis and phylogeny, as part of the process of adapting to the environment. The reported studies on compatibility and affordances leave open the question as to whether or not temporary/canonical and stable affordances, such as orientation and size, can be dissociated. Our paper aims to disentangle the role played by these different kinds of affordances during sentences comprehension.

In order to investigate this issue, we used a picture recognition task. Participants were presented with either an observation or an action verb (Look at / Grasp) followed by an object name. They had to decide whether the visual object following the sentence was the same as the one mentioned in the sentence. Objects graspable with either a precision or a power grip were presented in both a canonical and not-canonical orientation (see Figure 1). Following the embodied theory of language comprehension, we predict faster responses evoked by action rather than observation sentences, because the former rather than the latter sentences imply a physical interaction with the object involving the response effector. In addition, the task used invites to provide a judgement on the object, thus it probably renders the object characteristics particularly salient. Recent evidence (Bub, Masson & Cree, 2008) shows that a simple object noun activates both functional and volumetric affordances. According to the authors, functional affordances refer to using an object for its purpose, whereas volumetric ones refer to picking up the object.

Therefore, we predict that comprehending both an observation and an action sentence should lead to the formation of a "motor prototype" of the object, which encodes information on different affordances, and that the action sentences should have an advantage because both the noun and the verb involve an action hand.

---

Insert Figure 1 about here

---

Previous studies with the same paradigm have shown that we mentally simulate the shape and the orientation of an object implied by a verbal description. For example, reading the sentence “The ranger saw the eagle in the sky” led to faster recognition of an eagle with its wings stretched out rather than drawn in (Zwaan, Stanfield, & Yaxley, 2003). Similarly, “John put the pencil in the cup” led to faster recognition times than “John put the pencil in the drawer” if it was followed by a vertically rather than by a horizontally oriented picture of a pencil (Stanfield & Zwaan, 2001). These results support the simulation hypothesis. However, whereas the experiments by Zwaan and collaborators study the variability of affordances depending on the semantic context, the aim of our study is to assess whether observation and action sentences per se, without any reference to a specific context, activate affordances. We do not assume that affordances are independent of context, but rather that some affordances might be activated by default as it is highly likely they may provoke reactions across different contexts.

### 1.1. Experiment 1

The purpose of this behavioural study is to investigate whether or not language comprehension involves motor simulation and to test the role of stable and temporary affordances in such a simulation. More precisely, we intend to determine the effects of language comprehension on the representation of object action-related (and not just visual) properties (Glenberg & Robertson, 2000). In addition, we aim to determine whether we are equally sensitive both to stable and temporary affordances or whether there is some form of distinction between them.

We presented participants with either an observation or an action sentence followed by the object name (e.g., Look at / Grasp the brush). The photograph of an object followed the sentence. Participants had to decide whether the object presented in the photo was the same as the one mentioned in the sentence. In addition to sentences, we manipulated both the orientation and the kind of grip elicited by objects. Half of the photos represented everyday objects graspable with a precision grip (e.g., nail) and half with a whole-hand grip (e.g., brush). All objects were presented in both a canonical (i.e. the orientation typically linked to their functional use) and non-canonical orientation. For example, a candle (power grip) and a lit match (precision grip) were presented in either a canonical (i.e., with the graspable part located below) or non-canonical (i.e., with the graspable part located above) position. As previously described, with canonical orientation we referred in a broad sense to the orientation more typically linked to objects functional use. Namely, the objects we used did not have a large basis, thus their canonical orientation is unrelated to their typical static appearance but is rather related to how they appear while they are used. Consider for example a brush: when we pick it up it would have a horizontal orientation, but during use there would be a number of slightly different orientations, all more similar to the canonical orientation than to the non-canonical one.

If during sentence comprehension a simulation process takes place, then action sentences should produce faster RTs than observation sentences, as the first should more promptly evoke a motor response. In other words, action sentences should activate action information, and information on the different kinds of affordances, in a more powerful way than observation sentences, due to the fact that action sentences include a verb related to a manual action.

In addition, our hypothesis is that comprehending both kinds of sentences, leads to the potentiation of object stable and canonical affordances. In other words, while processing sentences on objects we may represent the object in the way we typically interact with it, thus building a sort of “motor prototype” of the object and activating its stable affordances (e.g., the typical grip it elicits) and its canonical affordances (e.g., its typical orientation). For example, we should represent a pen as an object graspable with a precision grip and with its affordance located below, given that we typically use pens in that position while writing (Barsalou, 1999). We predict faster response times when there is a correspondence between the properties of the motor prototype stored in memory and activated during sentence comprehension and the properties of the object represented in the photo. When information in memory and online visual information do not correspond, response time should be slower: the conflict between offline and online affordances should produce a response time cost.

Thus, if language activates a sort of “motor prototype” containing information on object canonical orientation, we predict faster responses for objects presented in a canonical rather than non-canonical position. In addition, if during sentence comprehension we build a simulation of the interaction with the object, we predict faster response times for objects graspable with a power grip than for objects grasped with a precision grip. Namely, in real life manipulating objects with a precision grip requires more time and effort than with a power grip, due to the higher complexity of the motor program implied by a precision grip (Ehrsson, Fagergren, Jonsson, Westling, Johansson, & Forssberg, 2000). In keeping with this, a number of behavioural studies provide evidence that precision grip object pictures are processed slower than power grip ones due to the activation of motor information (e.g., Borghi et al., 2007; Tucker & Ellis, 2001). Summing up, in contrast with predictions advanced by amodal theories of cognition (Fodor & Pylyshyn, 1988; Landauer & Dumais, 1997; Pylyshyn, 1973), if a motor simulation occurs during language processing, then:

a. we should find an advantage of action over observation sentences;

b. we should find evidence of the role played by grip (stable affordance);

c. we should find evidence of a conflict between information in memory and online visual information for orientation (temporary affordance);

d. with false trials, an interference effect should occur when the sentence is followed by the image of an object which activates the same kind of grip as that activated by the noun mentioned in the sentence;

e. with false trials, an interference effect should lead both to the disappearance of the advantage of canonical over not-canonical orientation and of action over observation sentences.

1.2. Experiment 2

In Experiment 1 we did not control the side (upper or lower) of the canonical affordance; in this way the objects we used might have a bias on location affordances. We designed Experiment 2 in order to control for this. First of all, we checked all objects in order to identify the location of the part affording canonical actions. Consider an object such as a brush. Different parts of it can be grasped in order to move it, but people would agree that the easiest and best way to use it is to grasp the handle. We might see a brush in different orientations, even though the most familiar orientation when we brush our hair is that of the brush with the handle located in the lower field. So we encoded the location of the handle for each object in typical use (e.g., for a brush it is in the lower part of the object). In order to make sure that there were no biases in the object structure, we selected an equal number of objects with canonical affordance in the upper and in the lower parts. Given that we presented each object both in the canonical and the non-canonical orientation, the standard affordance could be located either in the lower or in the upper field. By encoding both where the object affordance is located during the use of the object (canonical affordance), we could determine which aspects of response are due to information stored in memory and which stem from on-line information.

## 2. RESULTS

### 2.1. Experiment 1

5.86 % of the trials were removed as errors. Errors comprised all cases in which participants pressed the wrong key. RTs more than two standard deviations from each participant's grand mean for correct trials were excluded from this analysis. This trimming method led to the removal of 4.2% of the data. Mean RTs for correct response for true trials for each participant were submitted to a repeated measures ANOVA with sentence Verb (observation, action), object Orientation (canonical, non-canonical), and object Grip (power, precision) as within-subjects factors. Only significant results will be reported.

Given that the analysis of errors (excluding time-outs and errors in the catch-trials) revealed that there was no evidence of a speed accuracy trade-off, we focused on the RTs analysis. All main effects were significant. Action sentences were 16 ms faster than observation sentences, $F(1,29) = 5.27$, $Mse = 2883$, $p < .05$, canonical objects were 17 ms faster than non-canonical objects, $F(1, 29) = 9.55$, $Mse = 1766$, $p < .01$, and objects graspable with a power grip were processed 12 ms faster than objects graspable with a precision grip, *F* (1,29) = 4.91, *Mse* = 1908, *p* <.05. Given that the selected objects did not have a large basis, their canonical orientation appeared associated more to their use rather than to their typical appearance. Of primary interest was the interaction between Orientation and Grip (see Figure 2), *F* (1, 29) = 4.95, *Mse* = 2041, *p* <.04. When objects graspable with a power grip were in the canonical orientation, mean RTs were 30 ms faster than when they were in a non-canonical orientation (Newman Keuls, p <. 01). With objects graspable with a precision grip, the difference between canonical and non-canonical orientation was limited to 4 ms (Newman-Keuls, p =.65). This interaction suggests that the effect is not simply due to a perceptual factor, but it is rather modulated by the kind of Grip. The limitation of the advantage of upright orientation  to objects graspable with a power grip probably reveals a higher automaticity of power compared to precision grip. The marginally significant interaction between sentence Verb and Grip, *F* (1,29)= 3.68, *Mse* = 1996, p = .06, was also of interest, due to the fact that objects graspable with a power grip preceded by an action sentence evoked the fastest response times.

---

Insert Figure 2 about here

---

One further ANOVA was performed on correct responses to false trials. Consider that the factors of the ANOVA were different from those of true trials because there might be conflicting information between the sentence and the object appearing after it: for example, the sentence might refer to an object graspable with a power grip (e.g., brush), whereas the following object could be graspable either with a power or with a precision grip (e.g., spoon vs. pen). The factors manipulated within subjects were sentence Verb (observation, action), object Orientation (canonical, non-canonical), and Grip (same, different).The factors manipulated within items were sentence Verb and object Orientation, whereas the factor Grip was manipulated between items. No main effect was significant. Of primary interest was the interaction we found between Orientation and Grip, *F* (1,29) = 4.88, *Mse* = 2728, *p* <.04, due to the fact that RTs with objects in the canonical orientation graspable with the same grip as the objects mentioned in the sentence evoked the slowest response times (see Figure 3). This suggests that participants simulated a potential interaction with the object, and that this made for an increased difficulty in rejecting the objects in the canonical orientation graspable with the same hand posture. A possible objection is that one could predict facilitation instead of interference. We will handle this issue in the discussion.

---

Insert Figure 3 about here

---

2.2. Experiment 2

8 % of the trials were removed as errors. The data were trimmed with the same method used in Experiment 1. This led to the elimination of 5.6 % of the data. Correct mean RTs for true trials for each participant were submitted to a repeated measures ANOVA with the within subjects factors of Verb Type (observation, action), Object Grip (power, precision), Affordance Canonical Location (up, down), Field in which the affordance is located (upper, lower). A further ANOVA with items as random factor was performed, in which the factors of Verb Type, Object Grip, Field were manipulated within items, whereas the factor Affordance Canonical Location was manipulated between items. Given that the analysis of errors (excluding time-outs and errors with the catch-trials) revealed that there was no evidence of a speed accuracy trade-off, we focused on the RT analysis.

The main effects of Verb type and object Grip were significant. Action sentences were 16 ms faster than observation sentences, $F$ (1,18) = 5.35, $Mse$ = 4608, $p$ <.05, power objects were 21 ms faster than precision objects, $F$ (1,18) = 15.30, $Mse$ = 2267, $p$ <.01. This result was complemented by a significant interaction between the Verb type and the Affordance Canonical Location, $F$ (1,18) = 7.38, $Mse$ = 4260, $p$ <.01, suggesting that the advantage of action over observation sentences was mainly due to the objects with the stable Affordance Location in the upper position. We have no clear explanation for this effect: the frequency of handles is the same with objects with the Affordance Location in the upper and in the lower position.

Most importantly, the interaction between the canonical Affordance Location and the Field in which the affordance was located was significant, $F$ (1,18) = 5.72, *Mse* = 3130, $p$ <.03 (see Figure 4). Response times were faster in case of correspondence between the canonical affordance location (up, down), and the field in which it was presented (upper, lower). For example, when the ice cream was presented with the cone in the lower field, where it is typically grasped, RTs were faster than when it was located in the upper field. This interaction shows that language leads us to build an object prototype with the affordance located in a canonical position and demonstrates that information in memory may conflict with visual online information.

---

Insert Figure 4 about here

---

A further repeated measure ANOVA was performed on correct responses to false trials. The factors manipulated within subjects were sentence Verb (observation, action), Affordance Canonical Location (up, down), Field in which the affordance is located (upper, lower), and Grip (same, different).

No main effect was significant. However, the main effect of Grip almost reached significance, due to the fact that RTs with objects graspable with the same grip as the objects mentioned in the sentence were 19 ms slower than RTs of objects graspable with a different grip, $F$ (1,18) = 3.89, *Mse* = 6747, $p$ <.06. Of primary interest was the interaction we found between Grip and Verb, $F$ (1,18) = 9.15, *Mse* = 2507, $p$ <.01, due to the fact that the advantage of the action sentences over the observation sentences was limited to objects graspable with a different grip and that RTs with action sentences were slower with objects graspable with the same grip than with objects graspable with a different grip (see Figure 5).

This effect confirms the interference effect found in Experiment 1 with false trials.

Insert Figure 5 about here

## 3. DISCUSSION

### 3.1. Experiment 1

The predicted finding that action sentences would evoke faster response times than observation sentences suggests that a simulation process takes place. A recent study by Lindemann, Stenneken, Schie and Bekkering (2006) might explain our results. Participants were required to prepare grasping actions and to delay the execution of the motor action until a word appeared on the screen. On appearance of the word they were to perform a lexical decision or a semantic categorization task. The word could be either compatible or not compatible with the action goal (e.g. the word “mouth” was compatible with the action of grasping a cup and bringing it to the mouth, the word “eye” with the action of grasping a glass and bring it to the eye). In a control condition, during the preparation phase participants had to perform simple finger lifting movements. Response times were faster for words that are coherent with the action goal when participants were required to prepare a grasping action, but not a finger lifting movement. This suggests that semantic information is selected in accordance with the action intention of the actor. In our task participants had to judge whether the visual object was the same or different from the noun in the sentence; therefore the noun was probably the most salient word in the sentence. This can explain why a very similar simulation takes place with both action and observation sentences. However this simulation seems much more precise with action than with observation sentences (see also Experiment 2).

Our results have a further implication: in line with some very recent evidence (see for example Masson, Bub & Warren, 2008), they suggest that it is unlikely that reading an object name automatically activates a particular action regardless of the specific task used. Rather, the motor information activated by the object noun was modulated by the presented verb.

The marginally significant interaction between sentence and grip seems to show that, together with the canonical orientation, the simulation takes into account the grip (stable affordance) required by objects. In fact, action verbs led to a facilitation for power objects compared to precision ones, but there was no differential effects of the kind of sentence on object orientation. The advantage of power over precision objects with action sentences might be due to the fact that, compared to the precision grip, the power grip is less complex, more familiar, acquired earlier during development (Halverson, 1931) and it probably represents a more prototypical action of grasping compared to the precision grip. One could argue that, for objects like pens, precision grip is more adequate than power grip and engrained in one's experience with those objects. However, given that physically realising a power grip is less demanding than physically executing a precision grip, the very fact that objects graspable with a power grip are processed faster than those graspable with a precision grip suggests that during the task execution some kind of motor simulation has taken place.

The results show that processing sentences on objects activates a sort of "motor prototype" including stable and canonical affordances; that is, affordances related to both the kind of grip objects require and the canonical object orientation.

A first indication of this was the interference found in false trials provoked by objects in the canonical orientation graspable with the same grip as the objects presented in the sentence. This interference suggests that during language comprehension participants simulated a potential interaction with the mentioned object – and that this made for an increased difficulty in rejecting visual objects that possessed the orientation more likely to be grasped (canonical affordance) and that required the same kind of hand posture.

A possible objection is that one could predict facilitation instead of interference, because the microaffordances for objects that afford similar grip sizes overlap. Hence, they would evoke activity in the neurons responsible for perceiving the same affordances. However, we think the mechanism underlying the interference is straightforward. Namely, while comprehending the sentence the neurons responsible for a given kind of grip are activated. At the same time, the object canonical orientation is activated as well. It is well known that in AIP/F5 circuit, orientation, size and shape of visual objects (coded in AIP) are transformed into the appropriate motor schema (coded in F5) for acting upon them (Rizzolatti & Matelli, 2003). However in our case the activation of affordances does not refer to a visual object, that is it does not concern real time sensorimotor transformations, but it refers to a sentence about a manipulable object. Previous evidence (Borghi et al., 2004; Buccino et al, 2005 ; Glenberg & Kaschak, 2002; Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008; Scorolli & Borghi (2007) showed that words referring to graspable objects and verbs or sentences with a motor content, are able, as their referents, to activate the motor areas involved in action execution. Our data show that apparently both information on canonical orientation and size of the object can be automatically activated by the meaning of a noun.

Therefore a prototypical motor schema is activated that reflects the typical way (how and where) the object is grasped during its common use. However, in false trials, i.e. when the visual object is different from the noun, the prototypical motor schema automatically evoked by the noun does not correspond to the different object and has to be inhibited. The higher the similarity between the motor response elicited by the noun and that evoked by the object, the more difficult the inhibition is. This is the reason why the slowest responses are produced by objects requiring the same grip which were presented in the canonical orientation.

In addition, in false trials participants had to respond with the left rather than with the right hand. This obviously occurred even when target objects required the same grip as the objects mentioned in the sentence. Therefore, it is also possible that the "interference effect" could be the side effect of a specific motor program for the preferred hand. In other words, even though in our study we did not directly investigated the effects of manual preference on the response, the prototypical motor response for a given object might involve not only a particular set of affordances for the object, but also the hand that is most often used to execute the action, which would be, of course, the right hand for right handed people. The very fact that we find a selective interference for objects in their canonical orientation graspable with the same grip strongly suggests that the sentence evokes very precise postural information.

Two further results also support the idea that a “motor prototype” that included affordances related to both the kind of grip objects require and the canonical object orientation was activated.

On the one hand, objects graspable with a power grip were processed faster than objects graspable with a precision grip. The advantage of action over observation verbs with power grip objects confirms that the effect is not a visual, but a motor one, due to the fact that the power grip is easier to perform than the precision one. This result is consistent with the simulation hypothesis, given that in real life performing a precision grip is more complex and requires more time than performing a power grip. In addition, the interaction rules out the possibility that the effect reflects a bias in the material (i.e. that it is due, for example, to a higher typicality of power compared to precision graspable objects). On the other hand, the fact that objects oriented in the canonical position were processed faster than objects oriented in a non-canonical position, likely means that, at least for objects graspable with a power grip, when there was a mismatch between the “motor prototype” stored in memory and the visual information, processing required longer time. The effect of orientation might be due to the fact that participants automatically rotate reverse objects in order to simulate an interaction with them, and this produced longer reaction times. Nevertheless, at least one alternative account is viable. It is possible that the effect of orientation had simply to do with perception or facility of recognition of objects in the canonical orientation. Even if the facilitation limited to canonical power objects in true trials strongly suggests this was not the case because slower responses to objects in non-canonical orientations should be obtained with both types of objects, there might still be some bias in the materials. Experiment 2 was designed in order to control for these potential biases.

### 3.2. Experiment 2

The results confirm and extend the results of Experiment 1. First, the hypothesis that comprehending a sentence leads to a simulation process is confirmed, as the advantage of action over observation verbs suggests. Secondly, this experiment extends the results of Experiment 1 as it confirms that a motor prototype was created including information on object canonical orientation and grip. The existence of this prototype is suggested by the fact that RTs are longer when there is a mismatch between information in memory and online visual information. The interaction between Canonical Affordance Location and Field shows that seeing the object affordance (e.g., the handle) in a field different from the usual one implies a cost. This cost is clearly due to the activation of long term visuomotor associations between the agent and the object, that is, to the mismatch between information in memory and online information. In other words, this result can be an effect of "canonicalness", that is, an object presented in its most typical orientation will be more readily identified. Crucially, however, this "canonicalness" is not related to visual and perceptual aspects, but to the object functional use. For example, when we do not interact with a brush we do not typically see upright brushes – but when we use it to brush our hair or see somebody else combing his/her hair, then the canonical orientation would be more similar to the upright than to the reversed orientation.  In addition, the analyses of false trials strongly support the idea that language comprehension leads to the simulation of a motor prototype containing information related not only to canonical orientation but also to grip. Namely, the faster RTs obtained with the verb "grasp" as compared to the verb "look" are confined to cases in which the grip necessary to grasp the object mentioned in the sentence differs from the grip necessary to grasp the object presented visually. This finding is in line with recent evidence obtained by Masson, Bub and Warren (2008) which suggests that what they call "attention" verb (i.e. verbs that imply physically orienting to an object or its location but without physical contact) activate general functional information. This explains the results we obtained with observation verbs. Differently from observation verbs, our study reveals that action verbs do not activate general functional information but specific motor information, as indicated by the interaction between Sentence and Grip in false trials.

## 3.3. General discussion

### Synthesis

The data from the two experiments converge on the conclusion that comprehending sentences on objects activates a motor simulation. In both experiments we found that action sentences were processed faster than observation sentences and that objects graspable with a power grip evoked faster responses than objects graspable with a precision grip. Both these results are consistent with the simulation theory. In addition, in both experiments we found an orientation effect: in Experiment 1 canonical objects graspable with a power grip were the fastest, and in Experiment 2, in which the canonical affordance location was controlled, RTs were faster when the canonical affordance location and the field in which it was presented corresponded. The results on orientation suggest that a motor simulation takes place, but one could argue that they are due only to perceptual factors. However, the results on false trials show that a motor (not only visual) simulation takes place (for discussion on visual and motor simulation, see Jeannerod, 2006): in Experiment 1 we found a clear interference effect due to the fact that the picture of canonical objects graspable with the same grip as the object mentioned in the sentence evoke the slowest RTs. In Experiment 2 we found that an interference effect was modulated by the kind of sentence: thus, we found that with action sentences objects graspable with a different grip were faster than those graspable with the same grip; in addition, the advantage of the action sentences over the observation sentences was limited to objects graspable with a different grip. Importantly, the factor Grip was present in false trials of both Experiment 1 and Experiment 2, even if in Experiment 1 it was modulated by Orientation, and in Experiment 2 by Verb. In a very consistent way, in both experiments we found an interference effect when the required grip was the same.

Observation-action: neural basis

The simulation hypothesis accounts for the faster responses evoked by action rather than observation sentences in a manual task. This can be explained by the different neural circuitry activated by the two verbs. Recent studies on monkeys have shown that F5 area contains two varieties of visuomotor neurons: “canonical” and “mirror” (Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992; for a recent review see Rizzolatti & Craighero, 2004). Both canonical and mirror neurons discharge when macaques execute specific actions (for example, when they grasp an object with a precision or with a power grip). Canonical neurons also fire when the monkey simply observes an object, while mirror neurons fire when the monkey observes another monkey or an experimenter, but not a tool, performing a goal-directed action, such as, for example, grasping an object. Importantly, mirror neurons do not discharge when the object alone is presented. Recent studies have extended the discoveries on mirror neurons to humans (Buccino, Binkofski, Fink, Fadiga, Fogassi, Gallese, Seitz, Zilles, Rizzolatti, & Freund, 2001). In line with the simulation theory, studies on canonical and mirror neurons can help to explain the advantage of action over observation verbs. Namely, the observation verb probably activate only canonical neurons, that is neurons that fire in presence of the object alone, which probably express the past interactions with the object, or, in other words, the affordances of the object, whereas the action verb activates both canonical and mirror neurons (the latter fire when both the object and the action are presented). In line with this result, recent neurophysiological and behavioral data have provided evidence that the motor system is involved during action verb processing (Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005; Pulvermüller, Härle & Kummel, 2001; Scorolli & Borghi, 2007). At the same time, our study provides evidence that the simulation process occurs with both action and observation sentences, as demonstrated by the significant main effects of Grip (Experiment 1 and 2), of Orientation (Experiment 1), by the interaction between the canonical Affordance Location and the Field in which the affordance was located (Experiment 2), by the interaction between Orientation and Same vs. Different Grip in false trials (Experiment 1). The interaction between Grip and Verb in false trials in Experiment 2 and the marginally significant interaction between Grip and Verb in Experiment 1suggest, however, that the simulation is much more precise for action sentences than for observation ones.

Fine grained simulation

Our study suggests that the simulation triggered by sentences referring to graspable objects is quite fine-grained. In keeping with our results, recent studies favour the view according to which the simulation activated during language comprehension is quite detailed. For example, recent evidence shows that modifiers such as adverbs affect the meaning of verbs, therefore the meaning of sentences. Zwaan and Taylor (2006) showed that sensibility judgments for manual rotation sentences were made more quickly when the manual response to the sentence was in the same rotation direction as the manual action described by the sentence. This suggests that comprehension of manual rotation sentences produces motor resonance. In addition, they showed that motor resonance during sentence processing occurred relatively quickly and locally. By asking participants to read sentences like "Before /the / big race / the driver / took out / his key / and / started / the / car" while turning the knob one frame at a time, they found that the advantage in cases of congruency between actual turning direction and the motion implied by the sentence was localized in the verb region. In a recent paper, Taylor and Zwaan (in press) showed that, if a post-verbal adverb maintains focus on a matching action ("slowly" or "quickly"), motor resonance takes place. Interestingly, however, if the adverb shifts the focus to the agent (e.g., "obediently" or "eagerly"), a cessation of motor resonance occurs. In the same vein, De Vega, Robertson, Glenberg, Kaschak & Rinck (2004) asked participants to read short narratives describing an agent performing two actions that involved similar sensorimotor systems (e.g., chopping wood and painting a fence) or different ones (e.g., whistling a melody and painting a fence). The actions were described as simultaneous or successive by means of the temporal adverbs *while* and *after*, respectively. Comprehension was more difficult for sentences including the adverb *while* and actions involving the same sensorimotor system.

Motor prototype

The simulation evoked by sentences led to the formation of a sort of "motor prototype" of the object, which encodes the way to interact with the object in order to use it appropriately. Our results demonstrate that this "motor prototype" contains information on the canonical orientation of objects as well as on the kind of grip with which it is typically grasped. With object orientation, response times were longer when there was a mismatch between the motor prototype and the online visual information, i.e. when there was a mismatch between canonical and temporary information on object orientation. Importantly, our results show that affordances were activated even if they were not relevant to the task that required to recognize objects rather than to act with them. In addition, our results suggest that, compared to observation sentences, action sentences lead to the creation of a more clearly defined “motor prototype”. Compared to observation sentences, action sentences led to faster responses to permanent object features, such as size and canonical object orientation. In Experiment 1 action sentences evoked faster responses with objects graspable with a power rather than a precision grip, while in Experiment 2 responses to objects with canonical affordances in their upper part were faster.

Evidence congruent with the idea of the motor prototype is provided by a recent paper by Bub et al. (2008). The authors distinguish between two kinds of grasping gestures: those associated with functional use of an object (functional) and those used to pick up an object (volumetric). They found that responding to the color of an object was faster when the gesture learned to associate to the color matched either the functional or the volumetric gesture associated with the object. Their results indicate that motor information is automatically recruited both when viewing objects or making decisions about words. Importantly, however, their results favour a moderate version of the claim according to which motor information is “automatically” activated when seeing an object.

Automaticity?

As described in the introduction, there are a number of studies suggesting that visual stimuli automatically activate motor information. However, whether this activation is automatic or mediated by goals is still a matter of debate. Studies by Bub et al. (2008) suggest that behavioral effects are observed for objects only when the participants engage in a goal directed action that maps onto some of the actions afforded by the object. Similarly, Borghi et al. (2007) found a compatibility effect between a hand posture and an object graspable with it only when before the experiment participants engage in a goal directed action, even though they were not directly aware of the action goal. On the other hand, studies with eye-tracking paradigms provide evidence both of automatic activation of affordances that influences spoken word comprehension (e.g., Chambers, Tanenhaus & Magnuson, 2004) and of influence of transient semantic activation on affordances detection during spoken word recognition (e.g., Yee & Sedivy, 2006). In the same line, studies on the alien hand syndrome (for a recent review see Scepkowski and Cronin-Golomb, 2003) show that the upper limb of patients affected by this syndrome can perform rather complex motor activities outside of the voluntary control. The results of the present experiment do not directly address the issue of whether affordances activate brain areas responsible for goal-directed actions. However, they can provide some hints concerning this issue, as they are compatible with the view that reading words automatically activates a goal that selectively enhances a particular kind of motor response. Therefore, reading words that refer to a grasping action facilitate processing of an action more than of an observation sentence. Similarly, Lindemann et al. (2006) found that, after an intention to act has been formed, semantic processing of a word related to the action goal influences the motor response.

Kinds of affordances: neural basis

Though behavioural in nature, we believe our results are important for models of action and semantic knowledge organization in the brain. Overall, our results provide behavioral evidence in favor of the existence of different kinds of affordances and argue for the need to better disentangle the roles played by these different kinds of affordances in neural terms. The studies on compatibility and affordances described in the introduction leave open the question as to whether or not temporary/canonical and stable affordances, such as orientation and size, can be dissociated and referred to different cognitive and neural systems. Milner and Goodale (1995) have formulated the well known proposal that in the visual system there are two different mechanisms for the recognition of objects on the one hand and for object-directed actions on the other: an off-line mode that involves mainly the ventral stream and an on-line mode that depends mainly on the dorsal stream. Affordances are typically thought to be processed by the dorsal stream. However, if it turns out there are different kinds of affordances, they could well be sub-served by different neural pathways (Young, 2006). Temporary affordances, which refer to object properties that can vary depending on the context – for example the location of the handle – and are useful for accurate reaching and grasping, probably involve primarily the dorsal stream. However, these affordances lack information on object function, which can be given by the ventral stream and can influence action actualization. Therefore, canonical and permanent affordances, which depend on information stored in memory, may involve the ventral system. In accord with this view, various recent studies suggest that the distinction between the dorsal and the ventral stream as proposed by Milner and Goodale (1995) might be too rigid and dichotomic (Gallese, Craighero, Fadiga, & Fogassi, 1999; Derbyshire, Ellis & Tucker, 2006). For example, it has been proposed that the dorsal route can be distinguished into a pure dorsal-dorsal and a ventral-dorsal route (Gentilucci, 2003; Rizzolatti & Matelli, 2003). Further experimental and neuro-physiological studies are necessary to better understand the role played by different kinds of affordances.

## 4. PROCEDURE

### 4.1. Experiment 1

#### 4.1.1. Participants

Thirty students of the University of Bologna took part in the experiment. All had normal or corrected-to-normal vision, were right handed (self reported) native Italian speakers and were unaware of the purpose of the study. Participants were treated in accordance with the ethical guidelines of the Italian Psychological Association.

#### 4.1.2. Materials

Sixteen pictures of everyday objects were selected. We chose very simple manipulable objects, typically composed of two main parts (e.g., the handle and the opposite part). They all were elongated; that is, the vertical axis was longer than the horizontal axis. Half were graspable with a power grip (e.g., pear, brush), while half were graspable with a precision grip (e.g., cherry, pen). Each object was presented both in the canonical orientation (the orientation corresponding to its functional use) or in the non-canonical orientation (opposite to the canonical one).

We also created 3 different kind of imperative sentences, in which the verb was followed by a noun preceded by the determinative article. The sentences could include two critical verbs, an action verb (“Grasp”, *Prendi*) and an observation verb (“Look at”, *Guarda*). A third verb, “point” (*Indica*), was used as a catch-trial. Catch trials were introduced to induce participants to process not only the noun, but the whole sentence (see below). In Italian the verbs “grasp” and “look” are both transitive, and have the same length and the same frequency (to grasp, 512, to look at, 544, according to De Mauro, Mancini, Vedovelli & Voghera, 1993).

4.1.3. Procedure

The participants were tested individually. They sat in a dimly-lit room in front of a colour monitor. They were instructed to look at a fixation cross that was centrally displayed and remained in view for 1000 ms. Then a sentence appeared on the screen. After 400 ms the sentence was replaced by the target stimulus, the photograph of an object. We choose the presentation time in keeping with those of recent studies on sentence comprehension (Buccino et al., 2005). Namely, recent physiological studies are consistent with word representations in the brain that are activated as early as 100-200 ms after verbs onset and regardless of whether subjects focus their attention on the stimuli or not (Pulvermüller, 2001; Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006).

The timer started operating when the photograph appeared on the screen. The target stimulus remained on the screen until the participant responded, or until 2 s had passed. The participants were instructed to press the “9” key with the right index finger when the object in the picture was the same as that mentioned in the sentence and to press the “3” key with the left index finger when it was not. They were instructed, in cases where the sentence contained the verb “point” (catch-trial), to refrain from responding (see Figure 1). All participants were informed that their response times would be recorded and were invited to respond as quickly as possible while still maintaining accuracy. Participants received feedback after pressing the wrong key, after pressing a key after the catch-trial (“ERROR”), after taking 2000 ms to respond (“TOO SLOW”), or after correct responses (“CORRECT”).

One block of 48 practice trials was followed by a block of 384 experimental trials. The order of stimulus presentation was randomized and the factors of sentence Verb (observation, action, pointing), object Orientation (canonical, non-canonical), and object Grip (power, precision) were fully balanced. Of the 8 times each sentence was presented, 4 were followed by the object mentioned in the sentence in either the canonical or the non-canonical orientation (true trials) and 4 were followed by two different objects (false trials). For the false trials, the sentence was followed twice by an object requiring the same kind of grip as the object mentioned in the sentence (one with the canonical and one with the non-canonical orientation), and twice by an object requiring a different kind of grip (one with the canonical and one with the non-canonical orientation). For each trial, we recorded reaction times and errors.

4.2. Experiment 2

4.2.1. Participants

Nineteen students of the University of Bologna took part in the experiment. The selection procedure of participants was exactly the same as in Experiment 1.

4.2.2. Materials

Special care was taken in selecting the materials. Unlike Experiment 1, in which we used both artefacts and natural objects, in this experiment we used only artefacts, in order to eliminate any possible effect of intervening variables. Half of the objects were graspable with a power grip (e.g., brush) and half with a precision grip (e.g., pen). Most importantly, a pre-test was performed in order to evaluate objects' canonical orientation. We presented the pictures of the objects in the canonical orientation and asked a subgroup of 6 participants where the affording part was. In the end we selected 16 objects, half of them graspable with a power grip and half with a precision grip, and orthogonally half with the canonical affordance located in the upper, and half in the lower object part. Each object was presented with the canonical affordance in the upper or in the lower field. This allowed us to distinguish between the role played by temporary and canonical aspects/affordances pertaining to orientation. As in Experiment 1, the photographs were preceded by sentences referring to action (Grasp the upper / lower part of the candle) and to observation (Look at the upper / lower part of the candle), and by sentences working as a catch-trial (Point at the upper / lower part of the candle). Each sentence included the imperative verb, the object and an adverb indicating a direction (upward and downward). The two adverbs were respectively used in half of the sentences. We introduced the adverb in order to direct subjects' attention to a particular part of the object and to strengthen, in this way, the "typical" representation of it. Of the 8 times each sentence was presented, 4 were followed by an image of the object mentioned in the sentence with the canonical affordance in either the upper or lower field (true trials) and 4 by two different objects, one with power and one with precision grip (false trials).

4.2.3. Procedure

The procedure was exactly the same as in Experiment 1.

Literature references

References

Barsalou, L.W., 1999. Perceptual Symbol Systems. Behav. Brain Sci. 22, 577-609.

Barsalou, L.W., Simmons, W. K., Barbey, A.K., Wilson, C.D., 2003. Grounding conceptual knowledge in modality-specific systems. Trends Cogn. Sci. 7, 84-91.

Borghi, A.M., Glenberg, A.M., Kaschak, M.P., 2004. Putting words in perspective. Mem. Cogn. 32, 863-873.

Borghi, A.M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., Nicoletti, R., 2007. Are visual stimuli sufficient to evoke motor information? Studies with hand primes. Neurosci Lett. 411, 17-21.

Borreggine, K.L., Kaschak, M., 2006. The Action-sentence Compatibility Effect: Its all in the timing. Cogn. Sci. 30, 1097-1112.

Boulenger, V., Roy, A.C., Paulignan. Y., Deprez, V., Jeannerod, M., Nazir, T.A., 2006. Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. J. Cogn. Neurosci. 18, 1607-1615.

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

Buccino, G., Binkofski, F., Fink, G., Fadiga, R., Fogassi, L., Gallese, V. Seitz, R., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. Europ. J. Neurosi. 13, 400-404.

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 Res. 24, 355-63.

Chambers, C.G., Tanenhaus, M.K., & Magnuson, J.S., 2004. Actions and affordances in syntactic ambiguity resolution. J Exp Psychol Learn Mem Cogn. 30, 687-96.

Creem-Regehr, S. H., Dilda, V., Vicchrilli, A.E., Federer, F.,Ames N. Lee., J. N., 2007. The Influence of Complex Action Knowledge On Representations Of Novel Graspable Objects: Evidence From Functional Magnetic Resonance Imaging. J. Internat. Neuropsychol. Soc. 13, 1009–1020.

Corballis, M.C., 2000. Much ado about mirrors. Psychon. Bull. Rev. 7, 163-169.

De Mauro, T., Mancini, F., Vedovelli, M., & Voghera, M. 1993. Lessico di frequenza dell'italiano parlato. Milano: Etas.

Derbyshire, N., Ellis, R., Tucker, M., 2006. The potentiation of two components of the reach-to-grasp action during object categorisation in visual memory. Acta Psychol., 122, 74-98.

De Vega, M., Robertson, D.A., Glenberg, A.M., Kaschak, M.P. Rinck, M., 2004. On doing two things at once: Temporal constraints on actions in language comprehension. Mem. Cogn. 32, 1033–1043.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Exper. Brain Res. 91, 176-180.

Ehrsson, H.H., Fagergren, A., Jonsson, T., Westling, G., Johansson, R.S., Forssberg, H., 2000. Cortical activity in precision- versus power-grip tasks: An fMRI study. J. Neurophysiol. 83, 528-536.

Ellis, R., Tucker, M., 2000. Micro-affordance: The potentiation of components of action by seen objects. Br. J. Psychol. 91, 451-471.

Fodor, J.A., Pylyshyn, Z.W. 1988. Connectionism and cognitive architecture: a critical analysis. Cognition 28, 3-71.

Gallese, V., Lakoff, G., 2005. The brain´s concepts: the role of the sensory-motor system in conceptual knowledge. Cogn. Neuropsych. 21, 1-26.

Gallese, V., Craighero. L., Fadiga, L., Fogassi, L., 1999. Perception through action. Psyche, 5, 5-21.

Gentilucci, M., 2003. Object motor representation and language. Exp. Brain Res., 153, 260-265.

Gentilucci, M., Gangitano, M., 1998. Influence of automatic word reading on motor control. Europ. J. Neurosci. 10, 752-756.

Gibson, J.J., 1979, The ecological approach to visual perception, Houghton Mifflin, Boston.

Glenberg, A.M., 1997. What memory is for. Behav. Brain Sci. 20, 1-55.

Glenberg, A.M., Kaschak, M.P., 2002. Grounding language in action. Psychon. Bulletin Rev. 9, 558-565.

Glenberg, A.M., Robertson, D.A., 2000. Symbl grounding and meaning: a comparison of high dimensional and embodied theories of meaning. J. Mem. Lang. 43, 379-401.

Glover, S., Dixon, P., 2002. Semantics affect the planning but not control of grasping. Exp. Brain Res. 146, 383-387.

Glover, S., Rosenbaum, D.A., Graham, J., Dixon, P., 2004. Grasping the meaning of words. Exp. Brain Res. 154, 103-108.

Grafton, S.T., Fadiga, L., Arbib, M.A., Rizzolatti, G., 1997. Premotor cortex activation during observation and naming of familiar tools. Neuroimage 6, 231-236.

Halverson, H.M., 1931. An experimental study of prehension in infants by means of systematic cinema records, Genetic Psychology Monographs, 10, 107-286.

Hauk, O., Davis, M.H., Ford, M., Pulvermuller, F., Marslen-Wilson, W.D., 2006. The time course of visual word recognition as revealed by linear regression analysis of ERP data. NeuroImage 30, 1383-1400.

Huettig, F., Altmann, G.T.M., 2005. Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. Cognition 96, B23-B32.

Jeannerod, M., 2006, Motor cognition. What  actions tell the self, Oxford University Press, Oxford.

Kellenbach, M.L., Brett, M., Patterson, K., 2003. The Importance of Manipulability and Action in Tool Representation. J. Cogn. Neurosci. 15, 30-46.

Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., Wiley, C., in press. Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. Brain Lang..

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. Psychol. Rev. 104, 211-240.

Lindemann, O., Stenneken, P., Van Schie, H.T., Bekkering, H., 2006. Semantic activation in action planning. J. Exp. Psychol.: Hum. Percept. Perform. 32, 633-643.

Masson, M.E.J., Bub, E.M., Warren, C.M., 2008. Kicking Calculators: Contribution of Embodied Representations to Sentence Comprehension. J. Mem. Lang. 59, 256-265.

Milner, A.D., Goodale, M.A., 1995, The visual brain in action, Oxford Psychology Series, No. 27, Oxford University Press, Oxford.

Nazir, T.A., Boulenger, V., Roy, A., Silber, B., Jeannerod, M., & Paulignan, Y., 2008, Language-induced motor perturbations during the execution of a reaching movement, Q J Exp Psychol. 61, 933-943.

Palmer, S., Rosch, E., Chase, P., 1981. Canonical perspective and the perception of objects. In Attention and performance IX, J.Long & A. Baddeley, Eds., Erlbaum, Hillsdale.

Pecher, D., Zwaan, R.A., 2005, Eds., Grounding Cognition: The role of perception and action in memory, language, and thinking, Cambridge University Press, Cambridge.

Phillips, J.C., Ward, R., 2002. S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation. Visual Cogn. 9, 540-558.

Pylyshyn, Z.W., 1973. What the mind's eye tells the mind's brain: a critique of mental imagery. Psychol Bull. 80, 1-24.

Pulvermüller, F., 2001. Brain reflections of words and their meaning. Trends Cogn. Sci. 5, 517-524.

Pulvermüller, F., 2003, The neuroscience of language: On brain circuits of words and serial order, Cambridge University Press, Cambridge.

Pulvermüller, F., 2005. Brain mechanisms linking language and action. Nat. Rev. Neurosci. 6, 576-582.

Pulvermüller, F., Harle, M., Hummel, F., 2001. Walking or talking: behavioral and neurophysiological correlates of action verb processing. Brain Lang. 78, 143-168.

Riddoch, M., Humphreys, G., Hickman, M., Clift, J., Daly, A., Colin, J., 2006. I can see what you are doing: Action familiarity and affordance promote recovery from extinction. Cogn. Neuropsych. 23, 583-605.

Rizzolatti, G., Craighero, L., 2004. The mirror neuron system. Annu. Rev. Neurosci. 27, 169-192.

Rizzolatti, G., Matelli, M., 2003. Two different streams form the dorsal visual system: anatomy and functions. Exp. Brain Res., 153, 146-157.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., Buccino, G., 2008. Task related modulation of the motor system during language processing. Brain Lang. 105, 83-90.

Scepkowski, L.A., Cronin-Golomb, A., 2003. The alien hand: cases, categorizations, and anatomical correlates. Behav. Cogn. Neurosci. Rev. 2, 261-77.

Scorolli, C., Borghi, A.M., 2007. Sentence comprehension and action: Effector specific modulation of the motor system. Brain Res. 1130, 119-124.

Spivey, M., Geng, J., 2001. Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. Psychol. Res. 65, 235-241.

Stanfield, R.A., Zwaan, R.A., 2001. The effect of implied orientation derived from verbal context on picture recognition. Psychol. Sci. 12, 153-156.

Tanenhaus, M.K., Spivey-Knowlton, M. J., Eberhard, K. M., Sedivy, J. C., 1995. Integration of visual and linguistic information in spoken language comprehension. Science 268, 1632–1634.

Taylor, L.J., Zwaan, R.A., in press. Motor resonance and linguistic focus. Q. J. Exp. Psychol.

Tipper, S. P., Paul, M. A., Hayes, A.E., 2007. Vision-for-action: The effects of object property discrimination and action state on affordance compatibility effects. Psychon. Bull. Rev. 13, 493-498.

Tucker, M., Ellis, R., 1998. On the relations between seen objects and components of potential actions. J. Exper. Psychol.: Hum. Percept. Perform. 24, 3, 830-846.

Tucker, M., Ellis, R., 2004. Action priming by briefly presented objects. Acta Psychol. 116, 185-203.

Yee, E., Sedivy, J.C., 2006. Eye movements to pictures reveal transient semantic activation during spoken word recognition. J Exp Psychol Learn Mem Cogn. 32, 1-14.

Young, G., 2006. Are different affordances subserved by different neural pathways? Brain Cogn. 62,134-142.

Zwaan, R.A., 2004. The immersed experiencer: Toward an embodied theory of language comprehension. In Psychology of learning and motivation, vol.44, pp.35-62, B.H. Ross ed., Academic, New York.

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

Zwaan, R.A., Stanfield, R.A., Yaxley, R.H., 2003. Do language comprehenders routinely represent the shapes of objects? Psychol. Sci. 13, 168-171.

Figure Captions

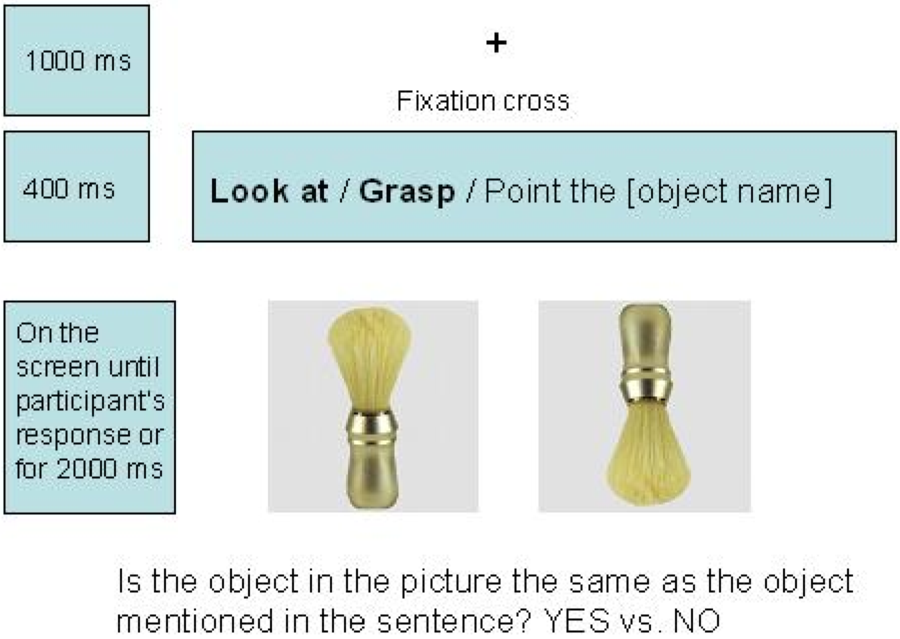
*Figure 1.* The experimental procedure.

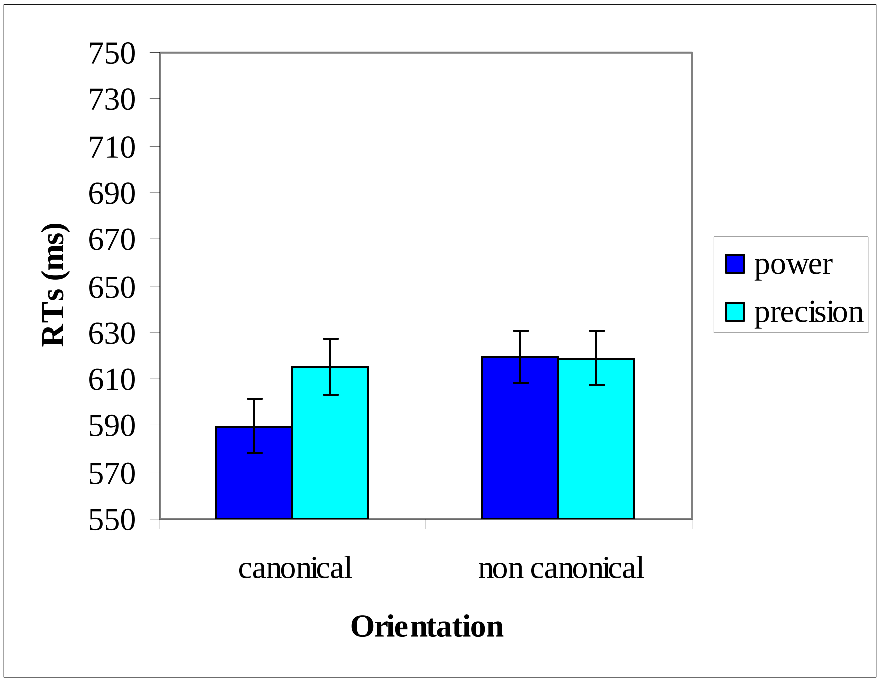
*Figure 2*. Experiment 1. True trials. The interaction between Orientation (canonical, non canonical) and Grip (power, precision).

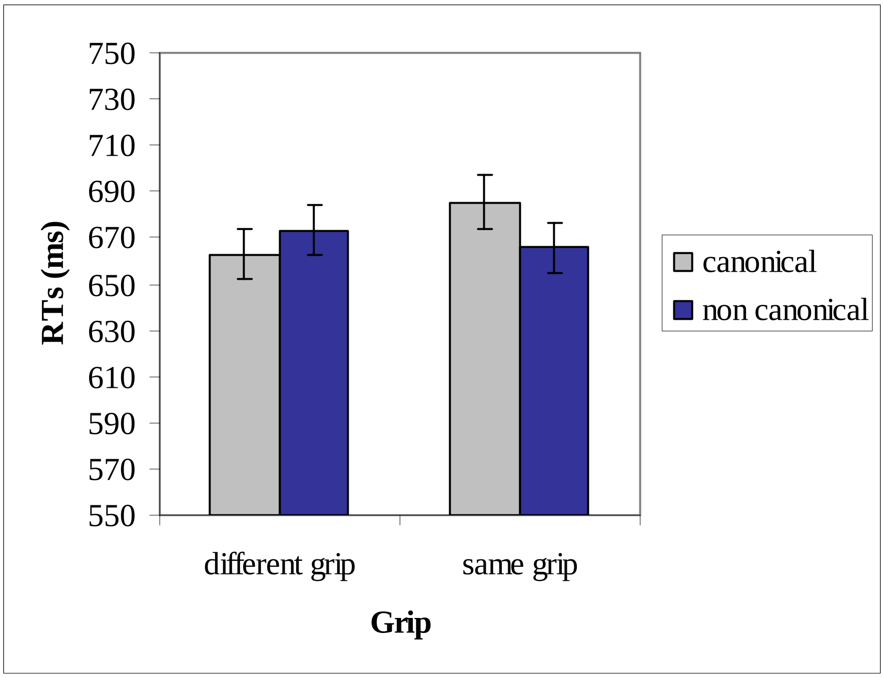
*Figure 3*. Experiment 1. False trials. The interaction between Orientation (canonical, non canonical) and Grip (same, different).

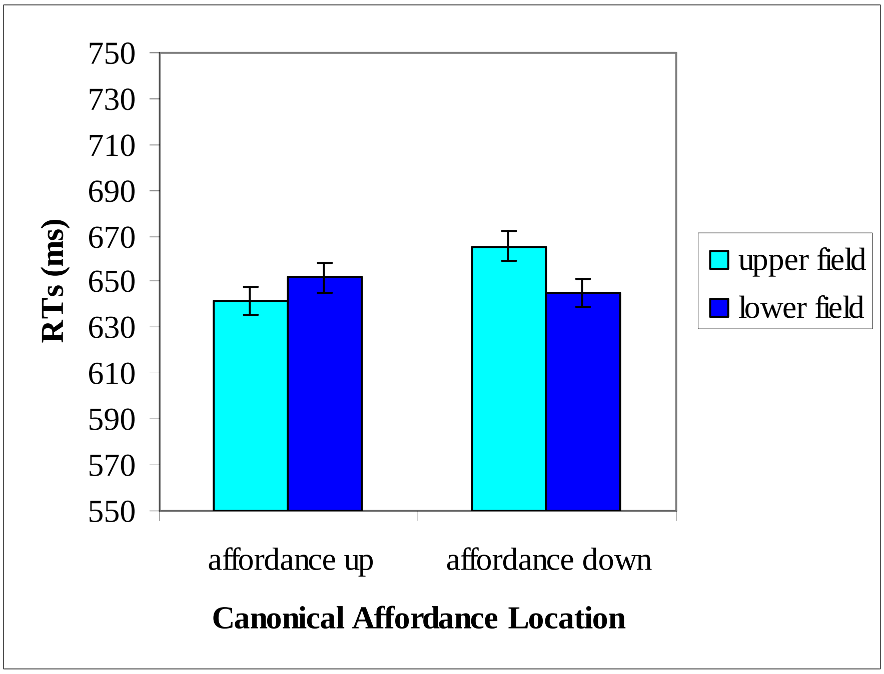
*Figure 4*. Experiment 2. True trials. The interaction between Affordance Canonical Location and Field.

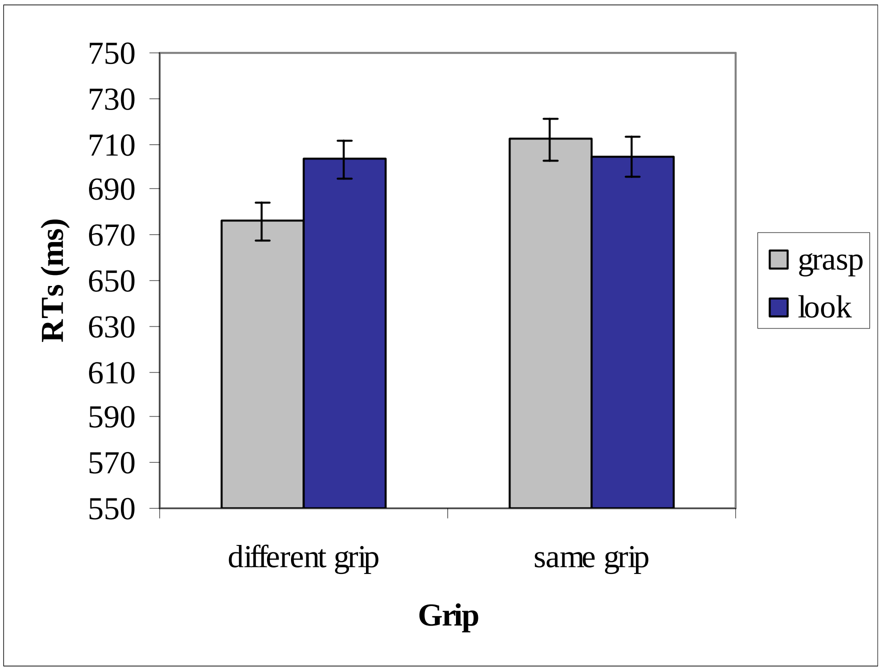
*Figure 5*. Experiment 2. False trials. The interaction between Verb and Grip.


1000 ms
+
Fixation cross
400 ms
**Look at** / **Grasp** / Point the [object name]
On the screen until participant's response or for 2000 ms
Is the object in the picture the same as the object mentioned in the sentence? YES vs. NO


750
730
710
690
670
650
630
610
590
570
550
RTs (ms)
power
precision
canonical
non canonical
Orientation


750
730
710
690
670
650
630
610
590
570
550
RTs (ms)
different grip
same grip
Grip
canonical
non canonical


750
730
710
690
670
650
630
610
590
570
550
RTs (ms)
upper field
lower field
affordance up
affordance down
Canonical Affordance Location



750
730
710
690
670
650
630
610
590
570
550
RTs (ms)
different grip
same grip
Grip
grasp
look