Language and embodiment

Claudia Scorolli and Anna Borghi

University of Bologna

claudia.scorolli2@unibo.it

annamaria.borghi@unibo.it

Classical view of concepts **2**

Embodied view of concepts **3**

The body, the context, and the current goals **4**

Neural bases of simulation: a brief outline **5**

Language and embodiment: behavioral, neuro-physiological and kinematics evidence **6**

1. Intrinsic properties: shape 6
2. Intrinsic properties: size 6
3. Intrinsic properties: color 7
4. Extrinsic properties: orientation 7
5. Modality 7
6. Perspective 8
7. Motion event 8
8. Action 9
9. Affordances 10
10. Effector and goal 10
11. The role of experience: ambiguous spatial word 11

Conclusion **12**

References **12**

Abstract

The paper focuses on the embodied view of cognition applied to language. First we discuss what we intend when we say that concepts are "embodied". Then we briefly explain the notion of simulation, addressing also its neuro-physiological basis. In the main part of the paper we will focus on concepts mediated by language, presenting behavioral and neuro-physiological evidence of the action/perception systems activation during words and sentences comprehension.

Key words

concepts, language, words, sentences, perceptual system, motor system, embodiment.

Classical view of concepts

The classical propositional view of concepts and meaning proposes that concepts are generated by abstract, arbitrary and amodal symbols (Collins & Loftus, 1975; Newell & Simon, 1976; Landauer & Dumais, 1997, Foltz, Kintsch & Landauer, 1998; Landauer, McNamara, Kintsch & Dennis, 2006). In this framework, the mind is a mechanism for syntactically manipulating symbols, such as an information processing device. Perception and action are considered as "low level" and peripheral processes, and low and high level processes are seen as reciprocally independent. In addition, perception and action are posited as separate spheres (Sternberg, 1969; Pylyshyn, 1999). Therefore it is not possible to envision action as having effects on perception, because the assumption is that the perceptual process takes place in the same way independently from the kind of motor response involved.

In this framework concepts are supposed to be "autonomous" from the body. They are represented in our mind in a propositional way, for example through list of properties, statements, frames, semantic networks (Fodor, 1998; Phylysin, 1973). According to this view a transduction process occurs, from the sensorimotor experience in the environment to the mind. The outcomes of this process are frozen representations of the world: in the course of the transduction every link with the body is lost. The ensuing representations are just arbitrarily linked to the world and do not have any modality specific feature: in this sense we could refer to them as *abstract symbols*. For example, the concept "dog" is associated with the amodal, propositional feature "it barks", rather than with the modal acoustic feeling of hearing a dog barking.

Accordingly mind is conceived of as the specific software evolved by humans for manipulating these abstract symbols. These symbols are organized in a stable-linguistic way, and they do not depend on the "hardware", that is on our body with its peculiar sensorimotor functioning.

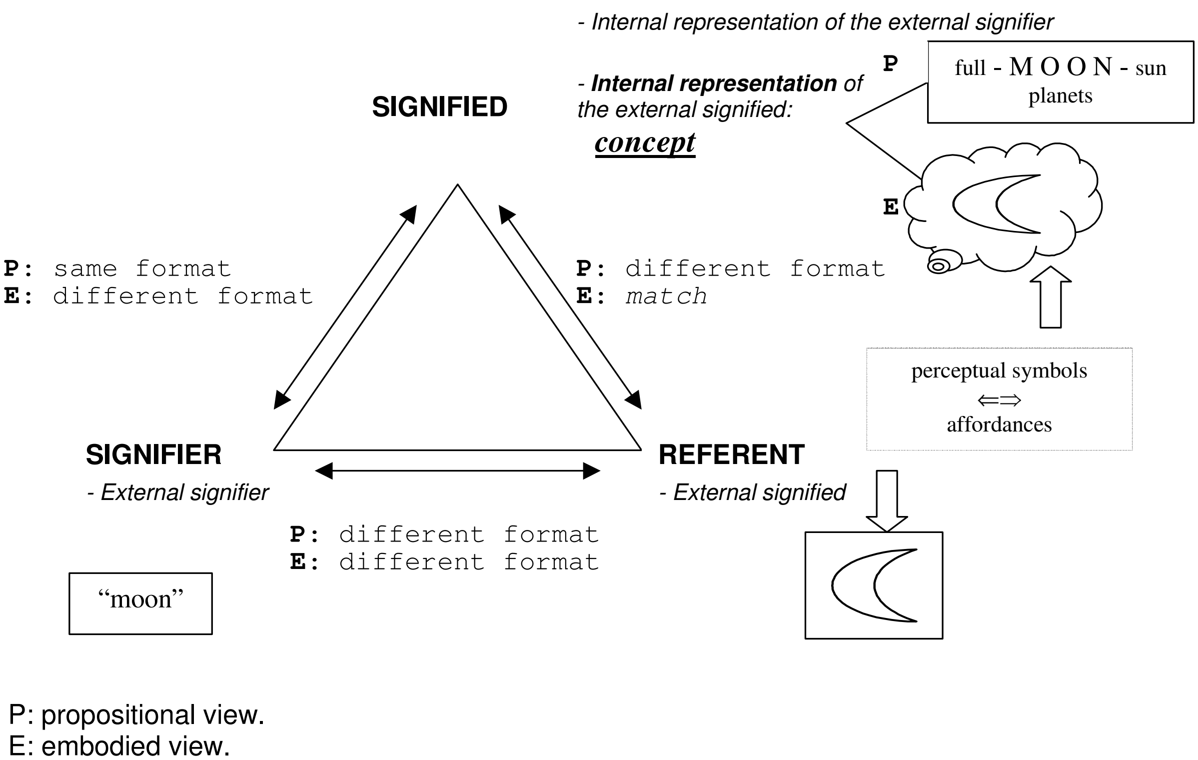
The consequence of this approach is the elaboration of models, for extracting and representing the meaning of words, based on statistical computations applied to a large corpus of existing texts. The underlying assumptions are that our knowledge is organized in a propositional way, and that the meaning of a concept/word depends on lexical co-occurrence and semantic relatedness. Examples of statistical models of semantic memory are the Hyperspace Analogue to Language (HAL, Burgess & Lund, 1997) and the Latent Semantic Analysis (LSA, Landauer & Dumais, 1997). In both the models word meanings is represented as vectors, detected in matrices (spaces with different dimensions) which describe the co-occurrence of terms in documents. That is: the meaning of a word is derived by its relations to other words and other abstract symbols. In this way it is possible *mathematically* / *spatially* calculating if two or more words/sentences are *equivalent*, namely if people represent them as semantically comparable or not. A low estimated parameter indicates that two words appear in different, *orthogonal,* contexts. The meanings of words are considered as fixed, so the understanding of a sentence would be pretty the same for everyone. LSA models outputs fit various experimental results: they fit human word sorting judgments and word-word lexical priming; they also successfully predict text learnability.

Nevertheless, there is much evidence that the predictions made by these models do not always match people understanding of sentences. For example, Glenberg & Robertson (2000) using LSA equivalent sentences, found that people actually distinguished between them depending on the perceptual characteristics of the objects. After a sentence like "Marissa forgot to bring her pillow on her camping trip", people judged more sensible and imaginable the sentence "*As a substitute for her pillow,* she filled up an old sweater with *leaves*" than "she filled up an old sweater with *water*", even though the words *"leaves"* and *"water"* are similarly *far* from *"pillow"* in terms of LSA norms. A pillow made by a sweater filled up with leaves is not usual, but it is afforded, so more sensible and imaginable than a pillow made by a sweater with water. Authors explain the results positing that "the meaning of words in sentences is emergent: meaning emerges from the mesh of affordances, learning history, and goals" (Glenberg & Robertson, 2000: 388).

Embodied view of concepts

Embodied view suggests that concepts are grounded in sensorimotor processes (Barsalou, 1999; Barsalou, 2008). They consist in the re-enactment of the same neural activation pattern running when we perceive their *referents* or when we interact with them (Gallese & Lakoff, 2005; Glenberg 1997). With "referent" we bear on the extra-linguistic reality, real or imaginary, to which the linguistic sign refers.

Revisiting Hjelmslev (1975) sign triad – that is an evolution of de Saussure's (1959) concept of sign – we could say that the propositional classical view of concepts assumes that the mental representations of the external signified has the same format and syntactical rules of the external signifier, intended as the linguistic sign. So language is mentally represented in terms of linguistic symbols and the relationship with the external referent is not taken into account.



Instead the embodied theory states that the format of concepts *matches* the format of their referents', i.e. our experience with/in the extra-linguistic reality to which they refer. In keeping with this view, the Indexical Hypothesis (Glenberg & Robertson, 1999) – that relates the general theory of embodied cognition to language comprehension – claims that language refers to objects and situations, or to the affordances of a situation (Kaschak & Glenberg, 2000) (for a straightforward explanation of "affordance", *see the third paragraph*).

The link between the mental representation of the signifier and the representation of the signified is arbitrary (for example, a “dog” is called “cane” in Italian). However, the internal representation of the referent is neither arbitrary nor abstract, but is rather grounded. Namely, in this view objects are represented in terms of perceptual symbols that are not arbitrarily linked but are rather analogically related to them (Barsalou, 1999). Perceptual symbols are multimodal, because they activate different motor and sensorial information tightly linked to the interaction with the world, pertaining vision, audition, taste, touch, motor action etc.

A notion useful to capture how language comprehension is conceived of for the embodied view is that of *simulation* (Gallese & Goldman, 1998). Simulating means that the same neural areas are implicated in perception and interaction with the objects, and in comprehending words that refer to them. For example, the word “glass” should reactivate the experiences of our previous interactions with glasses. So it leads to the activation of auditory, visual, and tactile information, for example the smoothness of a glass of wine, its sound banging into a dish, its shape and size, that surprisingly do affect the smell and the taste of the wine. The same word re-activates also proprioceptive and kinesthetic information, for example hand/arm feedback whereas bringing a glass to our mouth, as well as information on its affordance (for a straightforward explanation of “affordance”, *see the third paragraph).*

In the present work we do not discuss the embodiment of phonogical aspects of language (in the figure, “Internal representation of the external signifier”). For a critical discussion on this topic we refer to Liberman & Mattingly (1985), and to Fogassi & Ferrari (2007).

The body, the context, and the current goals

When we claim that cognition is “embodied” we don’t just mean that perception and action generically influence our knowledge, but that our knowledge specifically depends on having a *peculiar, unique body* and sensorimotor system. This implies that concepts cannot be static. The concept of a “chair”, for example, is not the same for a child and for an adult because of the different kinds of interaction a child or an adult could have with the referent, a real chair.

So the concept of a “chair”, conveyed by the object as by the written or listened word, in the case of an adult evokes the action of “sitting down on it”. The same concept evokes a different motor representation in an one-year-old child, such as for example the actions of “leaning on it, standing upright on it”. In Gibson’s terms (Gibson, 1979) we could claim that the *affordance* of a chair is not the same for an adult and for a child. The notion of affordance refers to the fact that objects in the environment offer to us as stimuli for acting, as if they would “invite” us to act upon them. Affordances are not fixed objects’ properties, but they are variable, as they depend on the interaction between objects’ features, our peculiar body and the surrounding environment. Glenberg (1997) remarks also the role of learning, pointing out that the meaning of a situation depends on affordances tuned on personal experiences of actions and learned cultural norms. The resulting set of available actions in turn depends on each individual’s present goals.

Importantly, according to the embodied view cognition is not only grounded in our body, but also *situated*, as it varies depending on the *context* and on the subject’s *goals*. So, for example, if we need to change a light bulb, the chair will no more afford some rest, but it will afford us a support for reaching the bulb. Therefore, our motor representations of the objects are guided by our current purposes and take into account both the constraints/possibilities of our body and the constraints/possibilities of our environment. In this perspective the subject, with his/her goals, is no more a passive spectator, and action is not simply the strict executive process that sequentially follows perception. The kind of motor response involved does have an effect on the perception of the present object/situation (Prinz, 1997; Hommel, Müsseler, Aschersleben e Prinz, 2001). Our conceptual knowledge is grounded and built on our action and interaction with the objects: *ago ergo cogito* (Glenberg, 1997).

Interestingly, the role of the body and of the context (the situatedness) in the representations is recently underlined also in social cognition. Also in this field the mental representations that underlie social behaviour were classically considered as abstract and stable. Recent evidence instead shows that also social representations, and the processes that underlie them, are adaptive and are modulated by the perceiver's current goals, communicative contexts, and bodily states (Smith, & Semin, 2007).

Neural bases of simulation: a brief outline

The neural substrate for the idea of simulation resides in the phenomenon of *motor resonance*. Recent neuro-physiological studies have led to important discoveries about the premotor cortex of the macaque monkey, the so called F5 area. This area contains two kinds of visuo-motor neurons: canonical and *mirror neurons* (Di Pellegrino, Fadiga, Fogassi, Gallese e Rizzolatti, 1992). Canonical neurons fire when a macaque executes specific actions, for example when it grasps an object with a precision grip, an all fingers grip or with a whole hand grip. They fire also when the macaque observes an object. Mirror neurons, instead, fire when the monkey observes a conspecific, or the experimenter, executing a goal-directed action, such as grasping a nut, but not when it observes just the nut.

Crucially to our aims, much evidence suggests that the homologue of F5 area in humans is the Broca area. In keeping with this, recently it has been demonstrated that the Broca area, traditionally known for its involvement in the production of language, contains a motor representation of the actions executed with the hand.

An fMRI (functional Magnetic Resonance Imaging: a technique to measure the hemodynamic response related to neural activity in the brain) study by Buccino, Binkofsky, Fink, Fadiga, Fogassi, Gallese, Seitz, Zilles, Rizzolatti and Freund (2001) showed that when subjects observe actions involving the mouth, the hand or the foot, different regions of the premotor cortex and of the Broca area are activated, depending on the different effector used for executing the action.

Symmetrically, also brain imaging and behavioral studies on language provide evidence of a somatotopic organization of the cortical areas. For example, Pulvermüller, Härle and Hummel (2001) have found topographic differences in the cerebral activity pattern generated by verbs relating to legs-, arms- and mouth-actions. Further evidence in favour of the tight link between language and action comes from a study by Buccino, Riggio, Melli, Binkofsky and Rizzolatti (2005) who performed both a TMS and a behavioural study. In the TMS (Transcranial Magnetic Stimulation: a noninvasive method to excite neurons in the brain) study they found that the motor evoked potentials (MEPs, that is a measure for the motor response) amplitude recorded from hand decreases during listening to hand action related sentences. They found a symmetrical motor response modulation on foot during listening to foot action related sentences. Consistently, in the behavioral study they showed that sentences describing actions with the hand or the foot activate the motor system in a specific way, that is: participants responded faster to hand related sentences if the response device was a pedal rather than a keyboard. A symmetrical modulation effect of language on motor system was found for foot related sentences.

Finally, as far as the tight link between the F5 area and the Broca area is concerned, it's worth to mention the proposal by Rizzolatti e Arbib (1998) about the important role that mirror neurons could have played in language evolution (see also Gentilucci & Corballis, 2006).

Language and embodiment: behavioral, neuro-physiological and kinematics evidence

An increasing body of evidence shows that language understanding implies a mental simulation (Gallese & Goldman, 1998; Zwaan, 2004). A heated debate within the embodied cognition community concerns whether the simulation activated during language comprehension is specific and detailed, or rather general. In the following part we will discuss studies performed using behavioral, neuro-physiological or kinematics methodologies, the results of which indicate that the simulation enacted by words is highly specific – that is, sensitive to the shape and orientation of the objects mentioned, to their motion direction, to the effector involved in the sentences etc.

We will first review studies pertaining objects intrinsic and extrinsic properties. With the term "intrinsic properties" we refer to objects invariant properties, like for example objects shape and size. Conversely, "extrinsic properties" are objects properties that depend on the observer or on the particular condition of observation, such as for example the current orientation of an object. In the second part, we will report evidence that highlight the role of different kinds of action and motor information for language processing.

1. Intrinsic properties: shape

Zwaan, Stanfield and Yaxley (2002) addressed if the simulation evoked during sentence comprehension is sensitive to subtle differences pertaining an intrinsic property, the shape. Participants were presented with sentences describing animals or objects in a different location, implying a different shape (e.g., [1] "He saw the lemon in a bowl" *vs*. [2] "He saw the lemon in the glass"). Their task consisted in deciding whether the picture represented a word mentioned in the sentence. The match condition led to an advantage in reaction times. For ruling out possible objections about the kind of task that could overtly require a comparison, authors designed a second experiment in which subjects had just to name the object/animal in the picture. In both experiments the results were straightforward: the response latency was lower when there was a match between the sentence and the picture (for example, when the sentence [2] was followed by the picture of a slice of lemon rather than of a whole lemon). The results suggest that while comprehending a sentence we automatically activate a perceptual representation, even if the current task doesn't claim for it. These results also show that sentence context has an important role in the building of these representations, that are dynamic and flexible.

2. Intrinsic properties: size

The specificity of the simulation for size was tested in different kinematics studies (Gentilucci & Gangitano, 1998; Glover & Dixon, 2002). The peculiarity of the kinematics method is that it allows detecting the activation of motor system during words processing. Glover & Dixon (2002), for example, asked subjects to reach and grasp objects on the surface of which either the word "Large" or "Small" were printed. The semantic meaning of the label shaped the early stages of both reaching and grasping movements; the semantic effect decreased over the course of the movement. A possible neurological explanation of the language effect on movement was ascribed to the closeness of language and motor planning centres, in the left hemisphere (Rizzolatti and Arbib, 1998).

3. Intrinsic properties: color

In an embodied view, the property of color is deeply different from typical multimodal properties, such as shape, because it is perceived by only one sense. The difference between theses two kinds of properties was well described by John Locke (1690/1975) who distinguished by *primary properties*, such as shape, size, and motion, that could be perceived by multiple senses, and *secondary properties*, such as color, taste, and smell, that are unimodal. He proposed that secondary properties could be represented less stably than primary ones.

Connell (2007) analysed whether implicit perceptual information about object color is accessed during sentence comprehension. Participants were presented with sentences that implied a specific color for the object described, as for example: "John looked at the steak on his plate". They had to decide if the picture showed after the sentence was mentioned in the sentence. The critical manipulation concerned the color of the picture: for example, either a brown or a red steak was shown. They found that perceptual information on color is activated during this task. However, participants were quicker when the object color implied by the sentence did not match the object picture color. The explanation they provide, consistently with the embodied view, is that accessing to shape, that is a stable property, is crucial for a recognition task. Thus if the color of the picture and the color of the object implied by the sentence do not match, there will be a minimal interference. Instead, if they match, the information on color is somehow difficult to ignore, even if color is a rather unstable property. This leads to a stronger interference on shape recognition.

4. Extrinsic properties: orientation

Stanfield and Zwaan (2001) demonstrated that we mentally represent the object orientation implied by a sentence in a figurative way. They showed participants a sentence suggesting a particular orientation of an object, for example horizontal or vertical (e.g., "He hammered the nail into the wall" *vs*. "He hammered the nail into the floor"). Then participants saw a picture showing the same object in an orientation that matched or not the orientation implied by the sentence. Responses were faster when the orientation suggested by the text matched the one of the picture. An amodal symbol system theory could possibly explain these results but it does not predict them.

5. Modality

In line with the view that perception, action and cognition are closely related, Pecher, Zeelenberg and Barsalou (2003) demonstrated that concepts activate multimodal information. They selected concept nouns and properties pertaining vision, motor action, audition, taste, touch and smell. Subjects were presented with a sentence like "A *lemon* can be *sour*". Their task consisted in judging if the sentence was true or false. Crucially, the task did not require to use mental imagery. Response times showed that switching modality, for example from a taste property (e.g.: *lemon* – *sour*) to an auditory property (e.g.: *leaves* – *rustling*), led to an increase in response times compared to the cases in which the modality remained constant. This demonstrates that concepts are multimodal rather than amodal. The only alternative explanation is that amodal symbols for the *same* modalities are more associated than amodal symbols for different modalities. This account was ruled out with a control experiment, in which they obtained analogous results (that is, slower response times when changing modality) using properties pairs much more associated than the pairs used in the first experiment. These findings clearly demonstrate that subjects simulate the content of the sentence, and that this mental representation activates a neural pattern in different modality specific domains. This explains why transferring processing from one brain system to another implies costs.

6. Perspective

Borghi, Glenberg and Kaschak (2004) demonstrated that the simulation we build during language comprehension is sensitive also to the perspective implied by the sentence. Participants read a sentence describing an object or a location from an inside (e.g., "You are eating in a restaurant"), an outside (e.g., "You are waiting outside a restaurant"), or a mixed (e.g., "You are walking toward and entering a restaurant") perspective. Then participants were presented with a concept-noun and they had to verify if the concept was or not a part of the location. For example, a "table" is a part typically found inside a restaurant, whereas a "sign" is typically found outside a restaurant. Responses were faster if the noun referred to an object more easily available in the perspective implied by the sentence. Interestingly, subjects also responded more quickly verifying that an object had a particular part if they were in the corresponding perspective and, within this perspective, for near than for far objects. Therefore, for example, they were faster if the *inside* sentence "You are eating in a restaurant" was followed by the *inside* part "table" than by the *outside* part "sign". In addition, the inside near part "table" was processed faster than the inside far part "kitchen". Results showed that the different perspectives suggested by the sentences control the accessibility of information, making available different conceptual knowledge.

In order to rule out a propositional explanation of the results, the authors computed the association degrees between sentences and parts, by using latent semantic analysis (LSA). The results were not explained by semantic associations between, say, *inside sentences* and *inside parts*: they are consistent with the idea that comprehension implies simulation. Finally, authors investigated perspective sentences that do not imply any action and that described an object in a particular orientation (e.g., "There is a doll upright in front of you"). Subject had to verify whether the noun presented after the sentence was a part of the object named in the sentence or not. The response device was a vertically oriented box. In order to provide a positive answer in the first part of the experiment participants had to move the hand upwards, and in the second part they had to move the hand downwards. Results showed faster responses when there was compatibility between the kind of response (yes-is-up vs. yes-is-down) and the location of the part (upper *vs.* lower). Crucially, these results were obtained using sentences that did not suggest any action.

7. Motion event

Two experiments by Kaschak, Madden, Therriault, Yaxley, Aveyard, Blanchard and Zwaan (2005) focus on *motion* direction, showing that the simulation we form is sensitive to the direction implied by the sentence suggested movement. Participants listened to sentences describing motion in four *directions*: away (e.g., "He rolled the bowling ball down the alley"), towards (e.g., "The dog was running towards you"), upwards (e.g., "The smoke rose into the sky"), and downwards (e.g., "The snow fell onto the ground"). Simultaneously they saw black-and-white motion perceptual stimuli: a clockwise and contraclockwise moving spiral picture (suggesting a motion away or towards participants' body); or an up or down moving horizontal stripes (suggesting a motion upwards or downwards participants' body). Subjects had to decide if sentences were sensible or not. Results showed that in the mismatch condition participants were faster than in the match condition.

In the second experiment participants were requested to make a grammaticality judgement on the sentence. The authors' purpose was to examine if the same interference effect was found with a task that did not emphasize semantic processing. Again, they found quicker response times for the mismatch condition. This data provide support for the claim that language comprehension is grounded in perception and action, and that the simulation activated by language is fairly specific.

The mismatch advantage (*interference effect*) found by Kaschak *et al*. (2005) is apparently in contradiction with previous evidence, showing faster response times when the sentence content matches the perceptual stimulus or the motor response. According to Kaschak *et al.* we need further empirical investigation in order to better understand these apparently contradictory results. However a possible explanation could be given by the interaction between two factors: the *temporal overlap* and the *integratability* (the degree in which the perceptual input could be integrated into the simulation activated by language). So the advantage of the mismatch condition could be due to the fact that visual perceptual stimuli engage the processing mechanisms needed to simulate the contemporarily listened sentences. The difficulty relies on the shared contents between the percept and the simulation of the sentence, and on the contemporaneous temporal overlap.

According to Kaschak *et al*., in the previously shown experiments there was an advantage of the match condition because the perceptual stimuli were easy to integrate with the sentence. For example, the sentence: "He saw the lemon in the glass" (Zwaan, Stanfield and Yaxley, 2002) was followed by the picture of a slice of lemon, and not by a black-and-white stimuli. The match advantage in the *integratability* condition is expected for a temporal overlap as for *sequentially* presented stimuli. Instead, when the visual stimulus and the sentence are processed sequentially and they are not easy to integrate there should be a null effect, because the stimulus is processed independently of the sentence.

The effect of language comprehension on visual representation of a *motor event* was also addressed by Zwaan, Madden, Yaxley and Aveyard (2004). Participants had to listen to sentences implying a movement toward or away from the body. Then they were presented with two pictures of a ball, differing only in size: the first one could be smaller or bigger than the second one, thus suggesting a movement toward or away from the observer. Subjects' task consisted of pressing two different keys to decide if the two pictures represented the same objects or not. In the match condition participants heard a sentence like "The shortstop hurled the softball to you" and then they saw a picture of a ball followed by a picture of a bigger ball. In the mismatch condition participants heard the same sentence and then a ball followed by a smaller ball. Results showed that in the match condition response times were faster than in the mismatch condition, suggesting that subjects activated a mental dynamic simulation of the sentence. More interestingly, these results were obtained with a task that did not involve in any way the content of the sentences. An amodal theory of cognition (Pylyshyn, 1986) can hardly account for these results.

8. Action

Further evidence for language grounding is provided by Glenberg and Kaschak (2002). They demonstrated that the simulation built during language comprehension is sensitive to directional aspects in action. Subjects were required to judge the sensibility of sentences moving the arm toward or away from the body. Half of the critical items referred to an action to perform by moving the arm toward the body, and the other half to a similar action done in the opposite direction. Critical items could be imperative sentences (e.g., "Put your finger under your nose" *vs*. "Put your finger under the faucet"), as well as sentences implying a concrete transfer (e.g., "Courtney handed you the notebook" *vs.* "You handed Courtney the notebook"), or sentences implying an abstract transfer (e.g., "Liz told you the story" *vs.* "You told Liz the story").

They found an *action-sentence compatibility effect* (ACE) in each of the three conditions. Thus, for example, responses to the sentence “Open the drawer” were faster if participants were required to perform a movement toward their body than away from their body; the opposite was true for a sentence like “Close the drawer”.

These results clearly support the idea that linguist meaning is grounded in bodily activity.

9. Affordances

The link between our knowledge and action was tested in a part verification task and in a sensibility judgment task by Borghi (2004). If concepts are represented as pattern of potential actions (Glenberg & Robertson, 2000) we would expect that different parts will be represented in a different way, depending on the more frequent action that we usually perform with the object. In other words, different objects parts can be good affordances depending on the situation at hand. So for example, in our representation of a “gas lighter” the *button* should be more salient than the *body* because *canonically* we use it for producing a spark. But the representation of the parts should change in relation to the requirements of the current situation. So for example, if we have just found a nice recipe in the book we were leafing through, the *body* gas lighter would become salient, allowing using the object as a bookmark. In the part verification task participants were faster when the object’s part word showed after the sentence was congruent with the part activated by the action suggested by the sentence. That is, “The child divided the orange” activates the mental simulation of the splitting action, and so the most salient part will be the “slice” rather than the “pulp” of the orange. The alternative propositional explanation, based on the semantic association between the verbs and the affording/nonaffording objects part, was ruled out replicating the results with controlled materials.

10. Effector and goal

Studies in different areas of neuroscience and cognitive science demonstrate that simulations formed during language comprehension are sensitive to the effectors implied by the verb or by the sentence.

Pulvermüller, Härle and Hummel (2001) investigated brain activity elicited by visually presented verbs that could be referred to movements of the arms (e.g., “to write”), of the legs (e.g., “to walk”) or of the face muscles (e.g., “to talk”). The behavioural part of the study consisted in a lexical decision task. In the physiological part they recorded Event Related Potentials (ERPs), that is a measure of the electrical activity produced by the brain in response to a sensory stimulus or associated with the execution of a motor, cognitive, or psycho-physiologic task. Behavioural results showed faster response times for face related verbs followed by arm related verbs and leg related verbs, supporting the idea that words semantic properties are reflected in the brain response they induce. Recorded ERPs revealed significant topographical differences 250 ms after stimulus appearance. Results seem to demonstrate that verbs that refer to actions performed using different effectors are processed in different ways in the brain.

Scorolli and Borghi (2007) also investigated the involvement of motor system in linguistic comprehension, using not single words but sentences composed by a verb and a concept noun. Verbs could refer to actions usually performed with the hands, the mouth or the feet. Subjects were requested to evaluate the sensibility of the sentences by pressing a pedal or saying ‘yes’ at the microphone. Response times showed that using the microphone they were faster with “mouth sentences” than with “hand sentences”. Using the pedal there was not a significant difference between “mouth sentences” and “hand sentences”; instead “foot sentences” were significantly faster than “hand sentences”.

This suggests that the same motor areas are recruited when a person understands action sentences or is actually performing the action. Importantly, this modulation occurred even with a task in which the information related to the involved effector was really irrelevant, such as the evaluation of the sensibility of sentences.

It's difficult to account for these results by means of abstract symbol theories of meaning. If words in these sentences are abstract, amodal and arbitrarily related to their referents, why did the effectors referred to by the sentence and used for responding influence the latencies of subjects' reactions?

In line with the previous studies, Borghi & Scorolli (submitted) demonstrate that the simulation activated by language is sensitive to the effector involved in the action expressed by the sentence and to the *specific* effector (right hand *vs.* left hand) used for responding, and that this sensitiveness seems to be modulated also by the goal implied by the sentence. Participants' task consisted in evaluating the sensibility of sentences regarding hand, mouth and foot actions (e.g., "Unwrap the candy" *vs.* "Eat the candy"; "Throw the ball" *vs.* "Kick the ball"). Participants responded by pressing two keys on the keyboard. The authors found a facilitation of sensible over non sensible sentences in right hand responses to hand and mouth sentences. This facilitation wasn't present in foot sentences. This finding suggests that the simulation evoked is quite detailed, as it is modulated both by the kind of effector the sentence refers to (hand *vs.* foot *vs.* mouth), and by the specific hand (dominant *vs.* non-dominant) the action expressed by the sentence typically involves.

The advantage of the dominant hand obtained with both hand and mouth sentences is particularly significant because it implies that people are sensitive both to the effector involved and to the goals expressed by the sentence. That is, mouth-related actions as "biting an apple" imply the simulation of the whole process of eating the apple, including bringing it to the mouth with the hand. On the contrary, the hand is typically not involved in foot related actions, such as "kicking a ball". The relevance of the goal is consistent with ideomotor theories (e.g., Prinz, 1997), that stress that actions are represented not only in terms of body movement but also in terms of the distal perceptual effects they aim to generate. The present data are also in keeping with Fogassi, Ferrari, Gesierich, Rozzi, Chersi, & Rizzolatti (2005) findings about a kind of mirror neurons that differentially codes a motor act according to the final goal of the action sequence in which the act is embedded. Finally, the results are convergent with evidence indicating that at the neural level hand and mouth actions activate contiguous regions, confirming the existence of a strict interrelationship between the effector hand and the effector mouth. This is in line with recent studies showing that language evolves from gestures and manual actions (e.g., Corballis, 2002; Arbib , 2005; Parisi, Borghi, Di Ferdinando & Tsiotas, 2005)

11. The role of experience: ambiguous spatial word

Embodied theories underline the role of the physical experiences in guiding concept understanding (Barsalou,1999; Glenberg, 1997; Wilson, 2002). Nonetheless, there is not much evidence on the role of experience in language comprehension. Most of the studies are restricted to paper-and-pencil tests (for example, concerning experience in motion, Boroditsky, 2000).

Alloway, Corley and Ramscar (2006) used a virtual environment for simulating an experience of motion. Virtual reality allowed them to directly test the embodied experiences on spatial perspective in order to investigate how ambiguous spatial terms are understood. Through virtual reality participants could experience either an ego-moving or an object-moving system. After having familiarized with the new environment, they were shown either an object moving linguistic prime (i.e., [1] "During the game, the green pillar is *in front of* the red pillar."), or a non-spatial question, unrelated to motion (i.e., [2] "During the game, most of the doors are closed.") They had to respond if the sentence was true or false. Results showed that participants were significantly influenced by the system of motion they represented. In fact, in the [1] case the prime overcame the embodied ego-moving schema of motion. On the contrary, in the [2] condition, in which the linguistic prime was unrelated to any system of motion, they were influenced by the ego-moving schema in the virtual environment, namely they responded to the target task coherently with the virtual suggested perspective.

Globally the results demonstrate that individual sensori-motor capabilities play an important role in guiding specific cognitive facilities. Focusing on language, it is crucial the finding that word meaning is not fixed, but influenced by our experience as well as by the linguistic context.

Conclusion

In our review we have shown that in the last ten years many studies have found support for the simulation theory, and have shown that the simulation we run during language comprehension is rather specific. Despite this huge amount of research, the amodal or propositional symbol system theory (Fodor, 1975; Pylyshyn, 1981) remains the dominant theory of knowledge representation. According to this theory the link between the internal symbols and the external referents is just an arbitrary one. It's hard to falsify this theory, because it can explain psychological phenomena. However, in many cases the explanation it can provide is just a *post-hoc* one.

Instead, according to the perceptual symbol system theory, the relationship between the symbols and their referent is not arbitrary, so a change in the referent will cause a change in the perceptual symbol (Barsalou, 1999b). The advantage of this theory is not in its explicatory power (even if this theory explains the same effects in a more parsimonious way), but rather in its predictive power.

References

- Alloway, T.P., Corley, M., & Ramscar, M. (2006). Seeing ahead: Experience and language in spatial perspective. *Memory and Cognition*. 34 (2), 380-386.
- Arbib, M.A. (2005). From Monkey-like Action Recognition to Human Language: An Evolutionary Framework for Neurolinguistics, *Behavioral and Brain Sciences*, 28(2):105-24.
- Barsalou L. W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences,* 22**,** 577-609.
- Barsalou, L.W. (1999b). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes,* 28, 61-80.
- Barsalou, L.W. (2008). Grounded cognition, *Annual Review of Psychology*, 59.
- Borghi, A.M. (2004). Objects concepts and action: Extracting affordances from objects' parts. *Acta Psychologica,* 115, 1, 69-96.
- Borghi, A., & Scorolli, C.(*submitted*). Sentence comprehension and hand motion simulation.
- Borghi, A.M., Glenberg, A.M., & Kaschak, M.P. (2004). Putting words in perspective. *Memory and cognition,*32, 863-873.
- Boroditsky, L. (2000). Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition,* 75 (1), 1-28.

- 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. *European Journal of Neuroscience,* 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 Research,* 24, 355-63.
- Burgess, C. and Lund, K., (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Process,* 12, 177–210.
- Collins, A.M. & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review,* 82, 407-428.
- Connell, L. (2007). Representing object colour in language comprehension. *Cognition,* 102, 476-485.
- Corballis, M.C. (2002). *From hand to mouth: The origin of language.* Princeton: Princeton University Press.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research,* 91, 176-180.
- Fogassi, L., & Ferrari, P.F. (2007). Mirror Neurons and the Evolution of Embodied Language. Current directions in psychological science, 16 (3), 136-141.
- 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.
- Foltz, P. W., Kintsch, W. & Landauer, T. K. (1998). The measurement of textual Coherence with Latent Semantic Analysis. *Discourse Processes*, 25, 285-307.
- Fodor, J. A. (1975). *The language of thought,* New York, Crowell.
- Fodor, J.A. (1998). *Concepts. Where cognitive science went wrong.* Oxford University Press.
- Gallese V. & Goldman A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences,* 12, 493-501.
- Gallese, V., Lakoff, G. (2005). The brain's concepts: The role of the sensorimotor system in conceptual knowledge. *Cognitive Neuropsychology, 21*, 455-479.
- Gentilucci, M., & Gangitano, M. (1998). Influence of automatic word reading on motor control. *European Journal of Neuroscience*, 10, 752-756.
- Gentilucci, M. & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, 30*,* 949-960.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin Company.
- Glenberg, A. M. (1997). What memory is for. *Behavioral & Brain Sciences*, 20(1), 1-55.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558-565
- Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, 28(1), 1-26.
- Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory & Language*, 43(3), 379-401.
- Glover, S., & Dixon, P. (2002). Semantics affects the planning but not the control of grasping. *Experimental Brain Research,* 146, 383–387.
- Hjelmslev, L. (1975), *Résumé of a theory of language.* Madison: Univ. of Wisconsin Press.

- Hommel,. B., Müsseler, J., Ascherslebenm, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioural and Brain Sciences,* 24, 849-878.
- Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory & Language*, 43(3), 508-529.
- Kaschak, M.P., Madden, C.J., Therriault, D.J., Yaxley, R.H., Aveyard, M., Blanchard A.A., & Zwaan, R.A. (2005). Perception of motion affects language processing, *Cognition,* 94 (3), B79–B89.
- 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.
- Landauer, T., McNamara, D.S., Dennis, S. & Kintsch, W. (Eds.) (2006). *LSA: A road to meaning*. Mahwah, NJ: Erlbaum.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- Locke, J. (1975). *An essay concerning human understanding.* In P.H. Nidditch (Ed.), Oxford: Clarendon Press. (Original work published 1690).
- Newell, A., & Simon, H.A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery,* 19, 113-126.
- Parisi, D., Borghi, A.M., Di Ferdinando, A., Tsiotas, G. (2005). Meaning and motor actions: Behavioral and Artificial Life evidence. *Behavioral and Brain Sciences,* 28, 35-36.
- Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, 14, 119-124.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129-154.
- Pulvermüller, F., Härle, M., & Hummel, F. (2001). Walking or talking? Behavioral and electrophysiological correlates of action verb processing. *Brain and language,* 78, 143-168.
- Pylyshyn, Z.W. (1973). What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery. *Psychological Bulletin*, 80, 1-24.
- Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, 87, 16-45.
- Pylyshyn Z.W. (1986). *Computation and cognition*, Cambridge MA: MIT Press.
- Pylyshyn, Z.W. (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22(3), 341-423.
- Rizzolatti, G., & Arbib, MA (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188-194.
- Saussure, F de. (1959). *Course in General Linguistics.* Eds. Charles Bally and Albert Sechehaye. Trans. Wade Baskin. New York: McGraw.
- Scorolli, C., & Borghi, A. (2007). Sentence comprehension and action: Effector specific modulation of the motor system. *Brain research*, 1130, 119-124.
- Smith E.R., & Semin G.R. (2007) Situated SocialCognition. Current Direction In Psychological Science, 16 (3). 132-135.
- Stanfield, R.A., & Zwaan, R.A. (2001) The effect of implied orientation derived from verbal context on picture recognition, *Psychological science,* 121, 153–156.
- Sternberg, S., 1969, The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica* , 30, 276–315.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review,* 9, 625-636.

- Zwaan, R.A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In: B.H. Ross, Editor, *The psychology of learning and motivation,* Vol. 43**,** Academic Press, New York, 35–62.
- Zwaan, R.A., Stanfield, R.A., & Yaxley R.H. (2002). Language comprehenders mentally represent the shapes of objects, *Psychological Science,* 13 (2), 168–171
- Zwaan, R.A., Madden, C.J., Yaxley, R.H., and Aveyard, M.E.. (2004). Moving words: dynamic mental representations in language comprehension, *Cognitive Science* 28, 611–619.