ANTHROPOLOGY & PHILOSOPHY

International Multidisciplinary Journal

Questo fascicolo è stato pubblicato con un contributo parziale
del Dipartimento di Studi Storico-Sociali e Filosofici
dell'Università di Siena

A&P

ANTHROPOLOGY & PHILOSOPHY
International Multidisciplinary Journal

Editorial Address:
Università di Siena - Dipartimento di Studi Storico-Sociali e Filosofici – Viale L. Cittadini, 33 - 52100 Arezzo (Italy) - Ph. +39-0575-9261 - Fax +39-0575-926312 - e-mail: bianca@unisi.it

A&P is published as one volume divided in 2 issues
Subscription Rates:
Institutional Subscription: 35 euro
Personal Subscription: 25 euro
To subscribe please write to:
A & P - Università di Siena - Dipartimento di Studi Storico, Sociali e Filosofici, Viale L. Cittadini, 33 - 52100 Arezzo (Italy)
Registrazione presso il Tribunale di Firenze n. 4822 del 6/8/1998

ANTHROPOLOGY & PHILOSOPHY
International Multidisciplinary Journal
Volume 9 (1-2) 2008

CONTENTS

Claudia Scorolli - Anna Maria Borghi
University of Bologna

LANGUAGE AND EMBODIMENT

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 and Loftus, 1975; Newell and Simon, 1976; Landauer and Dumais, 1997, Foltz, Kintsch and Landauer, 1998; Landauer, McNamara, Kintsch and 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; Pylyshyn, 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 ap-

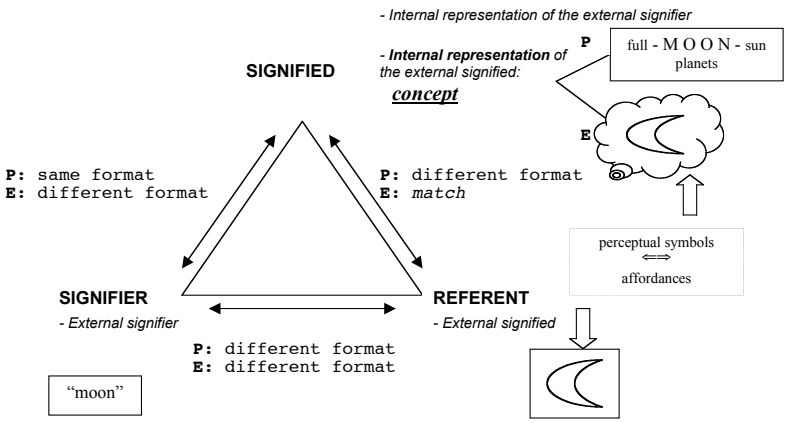
plied to a large corpus of existing texts. The underlying assumptions are that our knowledge is organized in a propositional way, and that the meaning of a concept/word depends on lexical co-occurrence and semantic relatedness. Examples of statistical models of semantic memory are the Hyperspace Analogue to Language (HAL, Burgess and Lund, 1997) and the Latent Semantic Analysis (LSA, Landauer and 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 and 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 and 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 and 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.



P: propositional view.
E: embodied view.

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 and 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 and 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 and 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 signi-

fier"). For a critical discussion on this topic we refer to Liberman and Mattingly (1985), and to Fogassi and 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 and 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 and 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 and 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 and Gangitano, 1998; Glover and Dixon, 2002). The peculiarity of the kinematics method is that it allows detecting the activation of motor system during words processing. Glover and 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 sug-

gested 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 and 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 and 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, and 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 and 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, be-

cause 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.M., 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., 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, 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, University of Wisconsin Press.

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

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, M.A. (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.M. (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 Social Cognition. *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 (Ed.) *The psychology of learning and motivation,* Vol. 43, Academic Press, New York, pp. 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., Aveyard, M.E. (2004). Moving words: dynamic mental representations in language comprehension. *Cognitive Science*, 28, 611-619.

Marco Cruciani
University of Trento

INTEREST AND MEANING

Abstract The problem of determining intended meaning is a key topic in the study of linguistic processes. This paper attempts to answer the question: how do agents involved in a linguistic controversy determine the intended meaning of a sentence? The main thesis of the paper is that the determination of meaning is driven by agents' situational interests. The process is analyzed in two phases (individual and contractual), and the thesis is respectively declined in two hypotheses. The first is that an agent's situational interest drives the individual choice of meaning for ambiguous sentences. It is argued in particular that formal semantics, the dictionary, context of use and domain knowledge are not sufficiently powerful to determine a unique meaning. From this it follows that an agent can *legitimately* choose a meaning (i.e. make a decision) given a set of acceptable meanings. This first proposal should impact on the problem of meaning under determination. The second hypothesis is that, in meaning negotiation processes, agents negotiate their own interests, and not directly the meaning. Meaning is then compatibly fixed on the agreement arising from the negotiation of interests. Focusing on cases of disputes concerning ambiguous clauses in employment contracts (collective agreements), I shall illustrate two linguistic controversies provoked by the same clause. The two controversies were resolved by determining two opposite meanings for the same expression. In the two situations, the two meanings were differently determined on the basis of different interests and two different negotiation processes. In this scenario I shall claim that the intended meaning can be regarded as an epiphenomenon of the negotiation of situational interests. This latter proposal should impact on the problem of determining the intended meaning in meaning negotiation processes.

Keywords Determining meaning, situational interest, decision, negotiation of interests, meaning as epiphenomenon.

Introduction

The paper seeks to answer the question: how is it possible to determine the intended meaning of a sentence within a linguistic interaction? Or better, how do the agents involved in a linguistic controversy determine the intended meaning of a sentence? It is not always possible to clarify all the linguistic and extra-linguistic rules relating to a particular context which enable us to establish the plausible meaning of a sentence. In my view, in some cases, when agents determine the meaning of a sentence, they resort to an individual decision, and then to negotiation between them.

In cases where ordinary tools are not sufficiently powerful to determine a unique meaning for a sentence or expression in a given situation, an agent chooses a meaning, from a given set of admissible meanings, which s/he believes to guarantee his/her own interest with respect to the specific situation (e.g. economic, social, moral, ideological, etc., interest). Once an individual choice has been made, the agent may be involved in two kinds of situation: in the first, the agent agrees with his/her interlocutors so that agreement is *de facto* determined. In the second, the agent does not agree with the other agents and a linguistic con-

troversy arises and a negotiation of meaning ensues. In my view, agents do not directly negotiate meaning, but rather their own situational interests. Thereafter they fix meaning compatibly with their contractual agreement.

The analysis of linguistic controversies arising from ambiguous clauses in contracts sheds clear light on real linguistic interactions, in which interactions are partially recorded, interests are sufficiently clarified, and meaning is truly important to the agents involved.

1. The Process and Two Hypotheses

The main thesis of the paper is that in the process of determination of the intended meaning of a natural language sentence there exists a relation between possible logical models of a sentence and extra-semantic interests of agents, or what they take to be their interests with respect to the situation. Essentially, agents' situational interests drive the determination of meaning. The process is analyzed in two phases: one individual and the other contractual, and the thesis is respectively declined in two hypotheses. These two dimensions are relevant because, as Clark puts it, "we cannot hope to understand language use without viewing it as joint action built on individual actions. The challenge is to explain how all these actions work" (Clark 1996: 4). Of course, Clark refers to communication, but I believe that this view is also valid for meaning negotiation.

The first hypothesis concerns the individual dimension of the process: in the case of a polysemous sentence, where meaning is important for an agent, s/he chooses a plausible interpretation on the basis of his/her own situational interest, making a choice from a given set of admissible meanings. In particular, given a specific situation where the same sentence, or expression, can have different meanings, even radically different ones, admitted by ordinary tools (i.e. formal semantics, the dictionary, context of use and domain knowledge), an agent can legitimately choose among them. An agent chooses an interpretation that s/he considers to satisfy his/her interest. In my view, it is very important that an agent can *legitimately* choose this interpretation, because ordinary tools leave open a semantic space in which an agent has no further linguistic constraints.

The second hypothesis concerns the contractual dimension of the process. In a linguistic controversy, the intended meaning reached after settlement of a dispute results from the negotiation of situational interests: that is, semantic agreement follows the agreement of interests. Agents negotiate their own interests and, after they have reached an agreement, they fix a meaning, among those admissible, which accords with their agreement. On this view, interests drive the process of determining meaning, and intended meaning is an epiphenomenon of the negotiation of situational interests.

I shall argue that formal semantics, use of the dictionary, support of the context of use and domain knowledge (e.g. criteria from the civil code) reduce the set of admissible meanings, but they are unable to determine a unique meaning. I shall show that each tool imposes certain constraints, and I shall represent the operation of those constraints by means of set-theoretical relation. I shall illustrate the tools by means of a schema in which each tool is represented by a level of a cone. Each level admits to a set of possible meanings for a sentence on the

basis of constraints specific to that level. I call this schema the "cone of language". In particular, the last level of the cone concerns the choice of a meaning among those selected by the previous levels. At this level, the set of possible meanings is partially ordered with respect to preferences which represent the agent's situational interests. At this point, the agent chooses a meaning and attempts to convince other agents that his/her meaning is plausible with respect to the situation.

In the last part of the paper, I shall illustrate two cases of linguistic controversy. In these two cases the same expression is debated and two different meanings are determined. I shall argue that a different semantic agreement is reached in the two situations, because different interests have been involved in the negotiation.

2. Some Basic Concepts: Meaning, Interest, Meaning Negotiation, Interest Negotiation

In this section I briefly introduce some concepts that I shall employ and discuss in the paper. I first consider the notion of meaning adopted in formal semantics, because this represents a stronger attempt to explain the semantic properties of language compared with other approaches (e.g. linguistics, conversation analysis, etc.). In semantics, the meaning of a sentence is fixed by its truth-conditions. Truth-conditions represent how the world would be if the sentence was true. If an agent knows the truth-conditions of a sentence, then s/he will also know its meaning. Hence an agent knows the meaning of a sentence even if s/he *de facto* does not know if the sentence is true (Wittgenstein, 1921; see Marconi, 1997; Casalegno and Marconi, 1992). For instance: an agent will know the meaning of the sentence "there is water on Uranus" if s/he knows the conditions in which water can be on Uranus: that is, for example, if s/he imagines some ice beneath the planet's surface. A central notion in formal semantics is that of model. Specifying a model for a language is a way to interpret it, to confer meaning on its expressions. Since sentences can be true or false with respect to the meaning that we assign to them, we say that a sentence is true or false with regard to a certain model (Casalegno and Marconi, 1992). According to the principle of the compositionality of meaning, the meaning of an atomic sentence is obtained by composition of the meanings of its constituents. From this it follows that if an agent knows the meanings of the words of a sentence, then s/he will also know the meaning of the sentence. Put otherwise: knowing the meaning of the words of a sentence is to know their semantic contribution to the meaning of the whole sentence. Similarly, the meaning of a compound sentence (such as a conjunction, a disjunction, or a conditional, etc.) functionally depends on the meanings of its atomic sentences (see Chierchia, 1992).

Another focus of the paper is the role performed by interest. According to Conte and Castelfranchi, an agent's interest can be represented as a state of world which favours the achievement of the agent's goal. An agent is interested in a certain state of the world p, if this implies another state q which corresponds to his/her goals (Conte and Castelfranchi, 1995). On this view, interest is a relation between one state of the world and another. In the case study, I shall report a

clause regulating the implementation of labour flexibility in two firms. A certain meaning of the clause (state *p*) implies a certain implementation of flexibility (state *q*) which favours the agent's goal (e.g. reducing labour costs).[1] An interest can be considered not only as a relation between states but also as a motivation for action which realizes a state. According to Latour, "as the name 'inter-esse' indicates, 'interests are what lie *in between* actors and their goals, thus creating a tension that will make actors only what, in their own eyes, helps them reach these goals amongst many possibilities" (Latour 1987: 108-109). On this view, an agent is interested in a certain meaning because s/he believes that it implies her goal.

How do agents negotiate their interests? As agents negotiate their interests, they negotiate their goals, and they reach agreement by giving and/or receiving resources (e.g. time, money, performances, objects) from other agents. In particular, agents have a negotiation problem whenever they have conflicting interests, or goals, and the realization of one agent's goal depends on the realization of the other agents' goals. Likewise, in the meaning negotiation view, according to Bouquet and Warglien, some agents have a semantic negotiation problem whenever they have "the problem of reaching an agreement on the meaning of an expression when (i) an agreement is valuable for all agents, but (ii) agents have conflicting preferences over which solution should be selected, so that every agreement implies that at least someone has to concede to some extent to other agent" (Bouquet and Warglien 2002: 2). In what follows, I shall show how agreement on interests selects out which is the intended meaning in a meaning negotiation.

3. Cone of Language

In the following subsections I shall argue that the support of formal semantics, use of the dictionary, context of use and domain knowledge are not able to determine a unique meaning for an expression. I shall argue that they are able to reduce possible meanings on the basis of some linguistic constraints, and I shall represent them by means of the 'cone of language', which is composed of four levels: formal semantics; use of the dictionary; context of use of an expression; and finally, the level of ordering preferences and decision making. The formal semantics level admits grammatically correct sentences. Semantics provides formal interpretations of sentences and reduces the potentially infinite set of meanings for a grammatically correct sentence[2]. The output from this level is a set of semantically admissible meanings for a sentence. The next level, that of the dictionary, provides interpretations for single terms and reduces the former set of meanings obtained by semantics. The next level, that of context of use and domain knowledge, provides knowledge which rules out some meanings with respect to the specific context and reduces the previous set of meanings obtained by dictionary. The last level concerns the individual choice. More precisely, it concerns the agent's preferences, and its output is a partially ordered set of meanings. At this point an agent chooses a meaning which accords with his/her preferences. Each level produces a set of meanings equal to or smaller than the previous one. We can write $M_n \geq M_{n+1}$, where M_n is the set of logical models ad-

mitted by the level n and M_{n+1} is the set of logical models admitted by the following level $n+1$.

For example, let us take the sentence: "the dog is dangerous". For brevity, let us assume that the word "dog" is given two interpretations by the dictionary: in the first, "dog" stands for canine mammal, and in the second it stands for pawl. Thus, in this simple case we have $M_2 = 2$ (at the level of the dictionary); but if, for instance, the sentence was written on a kennel, it would refer to a canine mammal. Hence, at the level of the context, the set of meanings becomes $M_3 = 1$ (in our example).

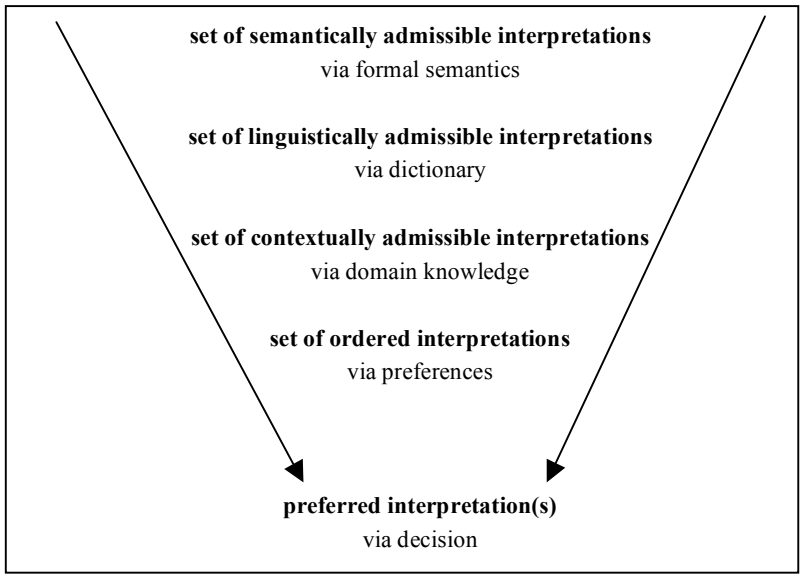


Fig. 1 Cone of language

3.1. Formal Semantics (or Standard View)

In this subsection I argue that formal semantics does not capture the meaning of single terms of natural language; with the consequence that it cannot determine the intended meaning of a sentence. Formal semantics (or model-theoretic semantics) is the approach to meaning which applies analytical tools used to study the semantic properties of formalized languages, like logic and mathematics, to natural language (Casalegno and Marconi, 1992). It can be regarded as the main paradigm in the analytic philosophy of language (Frixione, 1994)[3].

According to Bianchi, the main thesis in semantics is that the rules or conventions of a language are able to fix the meaning of every expression of that language. Semantics concerns itself with the meanings of linguistic expressions independently of the situation in which they are used, and it investigates the relation between linguistic expressions and objects in the world.

Semantics is:

- conventional: meaning depends on form of expression;
- truth-conditional: the meaning of a sentence identifies with its truth - conditions. And the meanings of words identifies with their contribution to the truth-conditions of the sentence in which they appear;
- compositional: the meaning of a compound sentence functionally de pends on the meanings of its components (Bianchi, 2003).

To provide a semantic interpretation of a natural language sentence, we must first translate the sentence into a formal language *L* expressive enough with respect to the universe of discourse, and we must then assign interpretations to its non-logical components. We assign interpretations by means of a function *I* that associates elements of *L* with objects in a domain *D* (the universe of discourse). The couple *M = (I, D)* is called 'model'; a semantic interpretation can be true or false with respect to a model.

Let us consider the English sentence (1) "there is a glass on a table" and translate it into the logical language of predicates. The syntax of *L* is composed of round brackets, universal and existential quantifiers (∀, ∃), an infinite set of variables (x_1, x_2, …, x_n), descriptive constants (*a* e *b*), predicative constants (G^1, T^1, O^2), truth-conditional connectives (~, ∧, ∨, →, ↔), and atomic formulas (*Ga*, *Gb*, *Ta*, *Tb*, *Oab*, *Oba*). Now we can write (1) as follows:

(2) $\exists x_1 \exists x_2 G(x_1) \wedge T(x_2) \wedge O(x_1, x_2)$.

Where *X* is a set of variables, *G* is the predicate "- is a glass", *T* is the predicate "- is a table" and *O* is the predicate "- is on -". Sentence (2) is a well-formed formula of *L* and corresponds to the syntactic structure of (1).

At this point we must provide a semantic interpretation of (2). Let *D* be a non-empty set and *I* a function that associates each individual constant of *L* with any elements in *D*. We consider two objects d_1 and d_2 belonging to *D* ($d_1, d_2 \in D$) such that:

- $d_1 \in I(G^1)$;
- $d_2 \in I(T^1)$;
- $d_1 \neq d_2$;
- $< d_1 , d_2 > \in I(O^2)$.

The conditions expressed above state that there are two objects which belong to the domain and respectively satisfy the predicate "- is a glass" and "- is a table"; they are different and satisfy the predicate "- is on -", that is, one is on the other one. But according to Marconi we do not yet have a real semantic interpretation, because the descriptive constants in (2) are not specifically explicated. In a standard semantic interpretation we specify only logical types of $I(G^1)$ and $I(T^1)$. Therefore we cannot choose among infinitive different models of a language (Marconi, 1997). Models can be infinite because infinite objects can satisfy the conditions in the above semantic interpretation. According to Frixione, formal semantics does not fix, for example, which sub-set of *D* the function *I* should associate with "glass". Similarly, it does not specify the difference between "glass" and "table". In essence, formal semantics explains the semantic

properties of a compound sentence through the semantic properties of its elements, but it says nothing about these elements, which are considered as given (Frixione, 1994). In this respect, formal semantics does not contribute to the problem of intended meanings in natural language because it does not contribute to resolving the problem of lexical meaning[4]. According to Thomason, the problems of semantics should be distinguished from those of lexicography. Formal semantics explains how different types of meanings are connected to different syntactic categories, but it does not explain how two expressions belonging to the same syntactic category differ with respect to their meanings. Formal semantics does not concern itself with which entities are correct or intended with respect to an interpretation of a language (Thomason, 1974). From this it follows that, in order to understand the intended meanings of natural language sentences, we must support formal semantics with respect to the lexicon. In my view the best candidate is the dictionary.

3.2. Dictionary

In this subsection I argue that the dictionary does not provide support powerful enough to individuate a unique meaning for a word, and therefore that it is not able to determine the intended meaning. This is so in two respects: in the case of polysemy, the dictionary does not provide a criterion with which to determine a unique meaning; but also in the case of a unique definition, the dictionary provides 'only' the conventional meaning, that is, it does not provide sufficient information with which to determine truth-conditions (*de facto* it underdetermines truth-conditions). Therefore in both cases it is unable to determine the intended meaning.

According to Putnam, it is the fact that we write dictionaries that founds the idea of semantics (Putnam, 1975). What is a dictionary? It is a list of words, each of which is followed by a definition of its meaning. A dictionary has at least three functions:

- f_1) it shows that a word exists and belongs to the vocabulary of the lan guage;
- f_2) it defines the meaning of a word by means of vocabulary;
- f_3) it provides some alternative interpretations, if necessary.

According to Bouquet, "dictionaries have two interesting properties:

- p_1) they provide a *publicly accessible* and *socially negotiated* list of accept able interpretations for a word;
- p_2) however, interpretations cannot *ipso facto* be equated with a list of shared meanings for the speakers of that language, as interpretations are (circularly) defined through other words, and do not contain the concept itself" (Bouquet 2007: 23).

On this view, a dictionary furnishes a network of socially accepted lexical relations which constrain concepts, but it does not contain them. Moreover, a dictionary should not contain factual knowledge like an encyclopaedia, but only linguistic definitions; hence, in principle, it cannot contain word references. Indeed, there is no clear demarcation line between a dictionary and an encyclopaedia, but I shall not deal with this topic (see Eco, 1997; Marconi, 1997; Quine,

1961). Here I (ideally) consider that a dictionary does not contain factual knowledge (e.g. see illustrated dictionary) and therefore cannot contain references for words. According to Bouquet, a dictionary can be viewed as containing socially reified lexical knowledge, and a linguistic community can be defined as a group of speakers who agree on a common dictionary (Bouquet, 2007). On this view, speakers can reach linguistic agreement independently of actual references of words. This issue shows a kind of intrinsic limitation of dictionaries in regard to determining references and concepts, but it throws light on the social dimension of certain linguistic constraints and on the possibility that intended meaning can be represented as semantic coordination with other speakers belonging to the same linguistic community (see Zanobini, 2006)[5]. However, I maintain that, in real situations, the dictionary provides a linguistic support which reduces possible interpretations in a linguistic community with respect to formal semantics.

In what follows, I report a case of an ambiguous clause and illustrate that the De Mauro dictionary is unable to provide a unique plausible interpretation of it. I have translated the clause from Italian into English, because the dispute occurred in Italy, but I believe it plausible that the analysis is equivalent. The clause runs as follows:

> In particular, it is understood that the employment relationship established with you will be resolved upon *cessation of the absence* of Miss Maria Rossi, and however not beyond 23 December 2005 [6].

This clause regulated the working conditions and hours of a recent graduate hired by an insurance company to substitute M.R. while she was on maternity leave. During the substitution period, M.R. resigned, and the employer terminated the graduate's employment on the grounds that the absence of M.R. had 'ceased'. The graduate argued that the clause referred to the absence of M.R. during maternity leave, not to her absence due to resignation.

The linguistic controversy arose with regard to the expression "cessation of the absence". We may use the dictionary to determine the meaning of the expression. I have translated the definitions from Italian into English. The Italian Demauro dictionary provides two possible interpretations for the word "cessation":

- a) termination;
- b) interruption, suspension.

It gives four possible interpretations for the word "absence":

- c) being away: *absence from work*;
- d) lack: *lack of air*, *lack of light*, *lack of gravity*;
- e) situation of uncertainty about a person's life subjected to a judicial decision: *declaration of absence*;
- f) temporary loss of consciousness: *in particular due to epilepsy, hysteria, intoxication* [7].

We can interpret the expression "cessation of the absence" through different combinations (compositionality) of the various interpretations of the single terms (e.g. the combinations *be* or *af*). The dictionary provides some socially accepted interpretations of words (function 3 and property 1) which constrain the admissible meanings of a sentence in which the words appear. On the other hand, the

dictionary is unable to determine a unique meaning for an expression, because it does not provide a criterion with which to choose one interpretation rather than another: they are on the same plane (although the dictionary can be used to check whether the meaning of an expression is admissible in a certain linguistic community). How can we decide which interpretation is plausible?

I first introduce the context of use (I shall consider it thoroughly in the next section). We can coherently individuate the expression "interruption of absence from work" among the others because it is compatible with the context under examination. For brevity, we suppose that it is the sole interpretation available from the dictionary. But if we compare it with the interpretations of the two agents, we see that it is compatible with both of them. It does not sufficiently specify the truth-conditions with respect to the real situation.

How can we discern which is the meaning in cases where we have several interpretations on the same plane; or, conversely, in cases where we have insufficient information from the dictionary? We can rely on a combination of domain knowledge and the context of use.

3.3. Context of use and domain knowledge

In this sub-section I argue that information concerning the context of use of an expression and domain knowledge do not suffice to determine a unique meaning for an expression, and that they are therefore unable to determine the intended meaning.

The philosophy of language considers two kinds of context: semantic and pragmatic. The semantic context represents relevant information through variables associated with the utterance: that is, it fixes the identity of speakers and interlocutors, the place, the time, etc. It contributes to determining literal meaning, and it is used in particular in cases of ellipsis, indexicality and ambiguity.[8] The pragmatic context is composed of a network of interlocutor beliefs, intentions and activities, and it contributes to determining the communicative intentions of speakers (Bianchi, 2003).

The pragmatic context can have pre-semantic and post-semantic uses (Perry, 1997). In the former case, the pragmatic context intervenes before the semantic context assigns an interpretation (e.g. in case of polysemy); in the latter, it intervenes after interpretation has been made to determine the actual communication ("speaker's meaning") with respect to the conventional meaning previously fixed by the semantic context (Bianchi, 2003). We need the pre-semantic and post-semantic uses of context because "the encoded meaning of the linguistic expressions underdetermines the proposition explicitly expressed by the utterance: [*conventional*] meaning underdetermines truth conditions. […] According to contextualism:

- i) the meaning of *any* sentence underdetermines its truth conditions - underdetermination becomes a *general property* of meaning;
- ii) the contextual factors that could be relevant for determining the truth conditions of a sentence cannot be specified in advance, and are not codified in the conventional meaning of the sentence" (Bianchi and Vassallo 2007: 78).

However, in the case examined here, some relevant pragmatic information is available in advance from domain knowledge concerning activities related to contracts. I believe it plausible that the goals, intentions and beliefs of agents must be at least compatible with contract-making rules (e.g. see civil code), otherwise a contract cannot be stipulated. I consider the context of use and domain knowledge to be collapsed together, and I refer to their intersection in accordance with what Bouquet and Giunchiglia (1995) called "context of work" (see also Bouquet 1998)[9]. On this view, the contextual information relevant to interpretation of a sentence in a particular context can be regarded as a subset of domain knowledge. I also consider that some information, such as the identities of the interlocutors, the place and time of the utterance, is trivially available in the context of contracts. I therefore assume that this kind of information belongs *latu sensu* to domain knowledge.

To clarify this point, I report two legal criteria (out of many) with which to disambiguate clauses of contracts stipulated under Italian law:

- *literal meaning* (art. 12 pre-laws civil code);
- *exegesis of common intentions of the parties* (art. 1362 c.c.).

These criteria are used to settle judicial and extra-judicial disputes. The former concerns the semantic context and pre-semantic uses of the pragmatic context; and the latter concerns post-semantic uses of the pragmatic context. Let us consider the first criterion: literal meaning. Suppose that we want to determine the conventional meaning of the sentence "the bank is wet". Before assigning an interpretation to the sentence, we must assign interpretations to its constituents: "bank" and "wet", and then compose the meanings of the single terms. But we cannot assign an interpretation to the single words because we must first know whether we are speaking about a river or a financial institute. In this sense, pre-semantic uses of context are based on specific information about the agent's activities related to the sentence. For example, "bank" may refer to "border of river" or "financial institute" and "wet" may refer to "damp" or "weak", respectively if an agent is going to fish or if s/he is talking to a board of directors. Hence, contextual information provides a criterion with which to decide which interpretations are plausible with respect to the situation. However, we have seen that in the case of "cessation of the absence" the literal meaning underdetermines the truth-conditions; hence both interpretations are compatible with the context. From this it follows that uses of the semantic and pre-semantic context are unable to indicate which interpretation is plausible.

The second criterion – "common intentions of the parties" – serves to determine what was the practical agreement between the agents, and subsequently, to determine the meaning compatible with the original agreement. However, some clauses/sentences admit different possible interpretations of 'common intention' which are legitimately defensible by the parties. In the case of "cessation of the absence", the common intention of the parties is to replace an employee (M.R.) with a temporary worker (graduate) during the former's absence due to maternity leave. It is not foreseen in the common intention that M.R. may resign and, in any case, the common intention does not capture certain aspects of the real situation. The disagreement arises because the company maintains that M.R.

is no longer an employee; on the other hand, the new graduate maintains that the maternity leave of M.R. has not ceased. Hence they must refer to the terms of the contract: "23 December" and not to the condition: "cessation of the absence". Essentially, in some cases even the contribution of the post-semantic context is not sufficiently powerful to determine a unique meaning for a sentence.

In cases of this kind, when there are no further linguistic tools available, how can we determine the intended meaning of a sentence? We must resort to a decision.

3.4. Individual choice

In this section I employ some basic notions from decision theory (Resnik, 1987; Myerson, 1991; Hansson, 1994) to describe the process of choice upon which decision-making is based. First of all, I wish to stress that, from an epistemological point of view, it is very important that an agent can *legitimately* choose an interpretation, among those admissible, because linguistic tools leave a semantic space open in which *de facto* every choice is legitimate.

Following decision theory, I maintain that agents' interests can be described with the language of preferences, and that a choice is made coherently with the agents' ordering preferences. Decision theorists have classified three kinds of choices: those made in conditions of certainty, in conditions of risk, and in conditions of uncertainty (ignorance). If a choice leads to a sure/assured state of affairs, then it has been made in conditions of certainty. If a choice leads to a possible state of affairs of which we can know the associated probability, then it has been made in conditions of risk. Finally, if a choice leads to a possible state of affairs, of which we cannot know its probability, then it has been made in conditions of uncertainty (or ignorance). This distinction is not clearly demarcated (Resnik, 1987). Nevertheless, here I assume that agents know that a certain meaning leads to a certain state of affairs. Indeed, the risk becomes real when the agents attempt to convince other agents.

However, the relation between interest and choice can be expressed through the language of preferences (Myerson, 1991; Resnik, 1987). Decision theory uses three comparative notions of preference: "better than" (>); "equal in value to" (≡); "at least good as" (≥). Using this language we can write, for instance, $(m_1 > m_2)_{Ag1}$, that is, agent Ag_1 prefers the meaning m_1 rather than the meaning m_2. Moreover, we can associate numerical values with the agent's preferences in order to represent their utility (relative importance). The convention adopted in decision theory is to assign higher numbers to higher options of ordering preferences. But numbers are not representative in themselves: what is representative is the proportion between them. However, here I shall take account only of ordering preferences, because they are sufficiently expressive for my purposes.

In our case, the set of admissible meanings M_3 represents the set of options which agents order with respect to their preferences. Decision theory assumes that a 'rational' agent can 'correctly' choose an option if the set of options is ordered in accordance with some formal properties[10]. Here, I consider only two essential properties: completeness and transitivity[11]. The formal property of completeness (for weak preference ≥) is defined for a relation and its domain:

- the relation ≥ is complete if and only if for any element A and B of its domain, either A ≥ B, or B≥A.

This property guarantees that an agent is able to compare between two options. The formal property of transitivity (for weak preference ≥) is defined as follows:

- the relation ≥ is transitive if and only if it holds for all elements A, B, and C, of its domain that if A≥B and B≥C, then A≥C.

This property guarantees that an agent is able coherently to compare among options. Indeed, "it is expected that the preferences that guide decisions are in many cases incapable of being represented by a complete preference relation. Nevertheless, in decision theory preferences completeness is usually accepted as a simplifying assumption […] although it is often a highly problematic assumption", as well as the assumption of transitivity (Hansson 1994: 17). But in our case, agents consider only a reduced set of all options because ordering involves only some possible interpretations: that is, the admissible meanings previously selected by linguistic tools. Therefore we can consider a partial ordering of preferences in which an agent 'coherently' makes a decision in accordance with the rule that states: "an alternative is (*uniquely*) *best* if and only if it is better than all the other alternatives. If there is a uniquely best alternative, choose it" (Hansson 1994: 19).

In synthesis, an agent legitimately chooses a meaning from a set of selected interpretations because it satisfies his/her interest. Now the agent must convince other agents that his/her meaning is plausible with regard to the specific situation.

4. Negotiation of interests and semantic agreement

In this section I illustrate two cases which concern two linguistic disputes provoked by the same clause in a collective agreement (national contract) stipulated by Labour Union and Confindustria (corresponding to the British CBI). I show that the two different negotiations of interests gave rise to two different intended meanings. I have translated the clause from Italian into English, as follows: "The parties agree on working hours, which apply also to groups of workers, with respect to flexibility regarding the *seasonality of products* […]. The parties further agree that, at company level, the modes and schedules of implementation will be reached through agreement with the local trade-union representatives" [12].

The dispute centred on the expression "seasonality of products". The two interpretations are: (a) 'time of the year' and (b) 'peaks in the market' respectively in both situations. The clause's meaning is important because of its impact on the criterion for implementing flexibility. In both cases, the respective interpretations are the same: in the former case the company sustains meaning *b*, and the local trade union sustains meaning *a*. Analogously, in the latter case, another company sustains meaning *b* and the same trade union sustains meaning *a*. In my view, it is very interesting that the same agent is involved in both situations and negotiates the same interpretations with different agents. In particular, I would stress that, in the two negotiations, different interests induced the same agent (the trade

union) to assume two different meanings. I call the two cases: "Draining pumps and seasonality of products" and "Fridges and seasonality of products".

4.1. Draining pumps and seasonality of products

In the case "Draining pumps and seasonality of products", the term 'flexibility' means that the company, during some periods of year, may require its employees to work a great deal of overtime. Overtime is required on Saturday or in addition to the regular eight hours. The company compensates overtime with paid rest days in others periods of year. Essentially, the clause regulates the times and ways in which the company can require overtime and compensate it with paid rest days.

The company is interested in managing working hours with discretionary power in order to save money, possibly on the basis of information unavailable to the local trade union (e.g. orders). In particular, the interest of the company is to be able to use overtime without paying the due wage supplements, and to distribute the cost of paid rest days in periods in accordance with its needs (its discretion). Moreover, the company is interested in being able to resort to overtime at any time of the year on the basis of market demand, and it is not interested in hiring new personnel or in paying overtime regularly. On the other hand, the trade union is interested in reducing (or avoiding) the use of overtime; and particularly if it is not regularly paid, in favouring the right to rest and to plan free time. It is also interested in inducing the company to hire new personnel and/or pay overtime regularly.

The company argues that overtime must be regulated with respect to peaks in the market: specifically, at any time of the year on the basis of market demand, the company can resort to unpaid overtime. The company cannot know peaks in the market in advance and thus cannot fix a specific period *a priori*. The trade union argues that overtime must be regulated with respect to the seasons of the year in which the products are most in demand, in particular the spring.

The two interpretations are both legitimate ($a, b \in M_3$), but different. At this point the parties attempt to reach a semantic agreement by negotiating their interests. In fact, the company and the trade union reached a semantic agreement where "seasonality of products" stood for "season of year when products are particularly in demand", in particular a 'positive season' is spring and a 'negative season' is autumn. They agreed that, in a positive season, the company can utilize not regularly paid overtime which in the negative season is recompensed with paid rest days (eighty percent of them on a collective basis and twenty percent on an individual one).

How did the agents reach the semantic agreement? How did the negotiation of interests work? The company obtained high discretionary power to utilize unpaid overtime in the positive season (from March to June) *de facto* independently of the peaks in the market and to collocate paid rest days in a period of year when it does not need labour, that is, during the negative season (from September to December). The company relinquished overtime throughout the year (except in the positive season) and discretionary power to distribute paid rest days during the negative season. The trade union obtained a reduction in unpaid over-

time (except in the positive season) and the right of employees to choose which days to use for paid rest during the negative season. Moreover, the trade union induced the company to hire new personnel or to pay overtime regularly (except in the positive season). The trade union relinquished checking overtime in the positive season. Finally, the trade union relinquished the possibility of distributing paid rest days throughout year, but only in the negative season.

The agent's interests were mediated with respect to the specific situation: each party gave up something in favour of the other party, and an agreement was finally reached. Thus meaning (*a*) was determined because it was compatible with the agreement reached.

4.2. Fridges and seasonality of products

In the case "Fridges and season of products", the term 'flexibility' means that the company can hire temporary workers and manage working hours and shifts in accordance with its needs.

The company is interested in hiring temporary workers on the basis of increased orders in any period of the year. The company is interested in managing temporary workers because of information unavailable to the trade union. The trade union is interested in reducing temporary work; in particular, the trade union is interested in restricting the use of temporary labour to only limited periods of the year. Moreover, the trade union is interested in reducing the use of temporary workers and in changing temporary jobs into salaried ones (on both permanent and fixed-term contracts).

The company claims that the use of temporary labour must be regulated in accordance with peaks in the market: that is, at any time of the year on the basis of market demand. The company cannot know peaks in the market in advance and thus cannot fix a specific period *a priori*. The trade union claims that the use of temporary labour must be regulated according to the seasons of the year in which products are most in demand, in particular the summer.

The two interpretations are both legitimate ($a, b \in M_3$), but different. At this point the parties attempt to reach an agreement by negotiating their interests. The company and the trade union reached a semantic agreement where "seasonality of products" stood for "peaks in the market". How did the negotiation of interests work in this case?

The company and the trade union reached an agreement in which the employer could use, in the case of peaks in the market throughout the year, temporary labour representing only ten percent of salaried labour (employees). Hence the company obtained high discretionary power throughout year, but only for a limited amount of workers. The trade union obtained a reduction in the use of temporary labour (ten percent of the workforce), but relinquished control over it. Finally, it lost bargaining power on new hirings.

The agent's interests were mediated with respect to the specific situation: each party gave up something in favour of the other party, and an agreement was finally reached. Thus meaning (*b*) was determined because it was compatible with the agreement.

5. Summary

We have seen that, in some linguistic controversies, ordinary linguistic tools are unable to grasp the meanings of natural language sentences and expressions. They reduce the possible meanings on the basis of different kinds of linguistic constraints, but they are unable to determine a unique meaning. Thus, an agent can legitimately choose an intended meaning from a set of admissible meanings previously selected by linguistic tools. S/he chooses a meaning which favours his/her interest.

We have then seen that agents involved in a meaning negotiation negotiate their situational interests, and not meaning directly. When they have agreed with respect to their interests, they subsequently fix a meaning compatible with the agreement. Hence, the intended meaning in a linguistic controversy can be considered as an epiphenomenon of the agreement on interests. This statement is plausible because, in the two negotiations, two different meanings have been determined for the same expression on the basis of two different negotiations of interests; even the same agent has determined two different intended meanings with regard to the two different situations. It is in this sense that situational interest drives the determination of the intended meaning.

Endnotes

[1] Intuitively, we could consider the state of the world *p* as the conditions which make a sentence true, and another state of the world *q* implied by *p* corresponding to the agent's goal.

[2] Sentences of natural language can be considered potentially as infinite (Chierchia, 1992), but our interest is to distinguish those that are semantically meaningful.

[3] Of course, Alfred Tarski's approach deals with artificial languages, but he recognized the importance of formal analysis in understanding natural language. Tarski (1944) claimed that linguistics is the most natural area to apply theoretical semantics.

[4] From this point of view, both extensional and intensional views are not sufficiently powerful to determine a unique meaning for a sentence. The former does not specify extensions of non-logical primitive symbols and the latter does not specify intensions of non-logical primitive symbols (Frixione, 1994).

[5] Note that in natural languages (e.g. Italian, English) there exist distinct dictionaries, and they may be different. Moreover, specific dictionaries exist in specific communities (e.g. web-users, farmers of an island community).

[6] Clause in Italian: "In particolare resta inteso che il rapporto di lavoro con lei instaurato si risolverà alla *cessazione dell'assenza* della signora Maria Rossi e comunque non oltre il 23 dicembre 2005".

[7] "Cessazione": a) *il porre fine, l'aver termine*; b) *interruzione, sospensione*. "Assenza": c) *l'essere assente, lontano*: assenza dal lavoro; d) *mancanza*: assenza d'aria, di luce, di gravità; e) *situazione d'incertezza sull'esistenza in vita di una persona scomparsa da almeno due anni, sancita mediante sentenza del giudice*:

dichiarazione d'assenza; f) *perdita momentanea della coscienza*, *spec. dovuta ad epilessia, ad isterismo, a forme d'intossicazione o ad eccessiva stanchezza*.

[8] Kaplan (1977) inaugurated a formal approach to semantic context in particular regard to indexical expressions.

[9] A context of work is composed of knowledge that an agent uses to resolve a problem. The question is: which knowledge pertains to the context of work? (Bouquet, 1998).

[10] Myerson (1991: 2-5) distinguishes between the rationality and intelligence of a decision-maker.

[11] In decision theory, the "basic properties that a rational decision-maker's preferences may be expected to satisfy can be presented as a list of axioms". They are: completeness, transitivity, relevance, monotonicity, continuity, objective substitution, strict objective substitution, subjective substitution, strict subjective substitution, interest and state neutrality (Myerson 1991: 9-12).

[12] Clause in Italian: "Le parti convengono, a titolo di *flessibilità sulla stagionalità dei prodotti* e per le attività di installazione e montaggio, sull'orario plurisettimanale, da realizzarsi anche per gruppi di lavoratori". […] "Le parti altresì concordano che, a livello aziendale, verranno convenute, tramite accordo, le modalità di attuazione oltre che i tempi di implementazione dell'orario settimanale di cui al presente punto con le rappresentanze sindacali unitarie e le organizzazioni sindacali territoriali".

Acknowledgments

I would like to thank Sonia Guglielminetti, Paolo Bouquet, Francesca Delogu, Roger Young, Marcello Frixione and Andrea Brighenti. This work was supported by the Network of Excellence: "Knowledge Web" – (FP6-507482).

References

Bianchi, C. (2003). *Pragmatica del linguaggio*. Roma-Bari, Laterza.

Bianchi, C., Vassallo, N. (2007). "Meaning, contexts and justification". In B. Kokinov, D.C. Richardson, T.R. Roth-Berghofer, L. Vieu (Eds.) *Modeling and Using Context*. Berlin Heidelberg, Springer-Verlag, pp. 69-81.

Bouquet, P. (1998). *Contesti e ragionamento contestuale. Il ruolo del contesto in una teoria della rappresentazione della conoscenza*. Pantograf, Genova.

Bouquet, P. (2007). "Context and ontologies in schema matching". In P. Bouquet *et al.* Proceedings of the workshop *Context and ontologies representation and reasoning*. Roskilde University, Roskilde DK.

Bouquet, P., Cruciani, M. (2004). "Il problema del ragionamento su azioni in intelligenza artificiale". In G. Ferrari (Ed.) *Profili mutlidisciplinari delle teorie dell'azione*. Franco Angeli, Milano, pp. 192-223.

Bouquet, P., Giunchiglia, F. (1995). Reasoning about theory adequacy. A new solution to the qualification problem. *Fundamenta Informaticae*, 23, 247-262.

Bouquet, P., Warglien, M. (2002). "Meaning negotiation: an invitation". In P. Bouquet (Ed.) *Meaning negotiation papers from the AAAI workshop*. Edmonton, AAAI Press.

Casalegno, P., Marconi, D. (1992). "Alle origini della semantica formale". In M. Santambrogio, *Introduzione alla filosofia analitica del linguaggio*. Roma-Bari, Laterza, pp. 42-87.

Clark, H.H. (1996). *Using Language*. Cambridge, Cambridge University Press.

Chierchia, G. (1992). "Logica e linguistica. Il contributo di Montague". In M. Santambrogio (Ed.) *Introduzione alla filosofia analitica del linguaggio*. Roma-Bari, Laterza, pp. 287-357.

Cruciani, M. (2008). "Meaning negotiation and situational interest". In Proceedings the 30th annual conference of cognitive science society. New York, Taylor and Francis.

Cruciani, M. (2007). "Situational interest and meaning". In Proceedings the 29th annual conference of cognitive science society. New York, Taylor and Francis.

Cruciani, M. (2007). "Interesse situazionale, decisione e significato". In G. Giardini and G. Pezzulo (Eds.) *Cognizione, complessità e cittadinanza*. Proceedings fourth Italian Conference on Cognitive Sciences, Roma, AISC.

Cruciani, M. (2007). "Negoziazione degli interessi e significato come epifenomeno". In G. Giardini and G. Pezzulo (Eds.) *Cognizione, complessità e cittadinanza*. Proceedings Fourth Italian Conference on Cognitive Sciences, Roma, AISC.

Cruciani, M. (2007). "Meaning and interests. Situational interest and meaning games". In P. Bouquet (Ed.) *Doctoral Consortium - Context* 2007. http://context-07.ruc.dk/Context2007DocCons.pdf.

Cruciani, M. (2006). "Interesse e significato". In A. Greco *et al*., *Scienze cognitive e robotica*. Proceedings Third Italian Conference on Cognitive Sciences, Genova, Erga.

Conte, R., Castelfranchi, C. (1995). *Cognitive and social action*. London, University College.

De Mauro, Italian dictionary on-line (http://www.demauroparavia.it/).

Eco, U. (1997). *Kant e l'ornitorinco*. Milano, Bompiani.

Frixione, M. (1994). *Logica, significato e intelligenza artificiale*. Milano, Franco Angeli.

Hansson, S.O. (1994). *Decision theory*. *A brief introduction*. Department of Philosophy and the History of Technology Royal Institute of Technology (KTH), Stockholm. http://www.infra.kth.se/~soh/decisiontheory.pdf.

Kaplan, D. (1978). On the logic of demonstratives. *Journal of Philosophical Logic*, 8, 81-98.

Latour, B. (1987). *Science in action. How to follow scientist and engineers through society*. Cambridge Massachusetts, Harvard University Press.

Marconi, D. (1997). *Lexical competence*. Massachusetts, MIT Press.

Marconi, D. (1999). *La filosofia del linguaggio*. *Da Frege ai giorni nostri*. Torino, UTET.

Montague, R. (1974). *Formal philosophy*. New Haven e London, Yale University Press.

Myerson, R. B. (1991). *Game theory*: *Analysis of conflict*. Cambridge (MA), Harvard University Press.

Perry, J. (1997). "Indexical and demonstratives". In B. Hale, C. Wright (Eds.) *A Companion of philosophy of language*. Oxford, Blackwell, pp. 586-612.

Putnam, H. (1975). *Mind, language and reality*. Cambridge, Cambridge Univesity Press.

Resnik, M.D. (1987). *Choices*. Minnesota, Minnesota University Press.

Tarski, A. (1944). The semantic conception of truth and the foundation of semantics. *Philosophical and Phenomenological Research*, 4, 341-76.

Thomason, R. (1974). *Introduction to Montague*. In Montague (1974) New Haven e London, Yale University Press.

Quine, W.V.O. (1961). Two dogmas of empiricism. *The Philosophical Review* 60 (1951): 20-43. Reprinted in W.V.O. Quine, *From a logical point of view* (Harvard University Press, 1953; second, revised, edition 1961).

Wittgenstein, L. (1921). *Tractatus logico-philosophicus*. Oxford, Basil Blackwell 1961.

Zanobini, S. (2006). *Semantic coordination*. *The model and an application*. PhD dissertation, Dept. of Information and Communication Technology, University of Trento, Italy.

Tiziana Giudice
University of Bari

METAPHOR BETWEEN EMBODIMENT AND IMAGINATIVE PROCESSES

Abstract In this paper I will analyse the relationship between metaphor and imagination. This issue has been recently studied by cognitive linguists who appreciate its importance, while other semantic perspectives neglect it. I will analyse the thesis which affirms that metaphors are based on cognitive components which are not logical-propositional but imaginative: the "image schemata" are recurrent models of corporeal experiences, centres of knowledge organization which structure – in a non-propositional form – an amount of salient information. This information emerges from the sensorimotor activities such as: manipulation of objects, spatial orientation and focusing perceptive attention on different goals. According to this theory, imagination is the drive of the entire cognitive experience and, through metaphorical projection, it activates and coordinates the schemata from which the very linguistic activity originates. Starting with empirical studies about the cognitive and neurobiological plausibility of image schemata, I will assert the function of imagination as the *trait d'union* between experience and cognitive use of the metaphor, through the use of sensory and synesthetic metaphors as mediators among perceptive, physical and linguistic dimensions. I hope future research will concentrate on the relationship between synesthesia and image schemata because of their strong connection with human perceptive, cognitive and neural structures.

Key words Metaphor, embodiment, image schema, synesthesia, simulation.

1. Introduction

In this essay, I will analyse a particular aspect of the metaphorical device, that is its relationship with the imagination. The metaphor, from the perspective of Cognitive Linguistics which concerns us here, is a strategy adopted by the mind to conceptualize reality (Lakoff and Johnson, 1980). Faced with the inadequacy of literal language in dealing with the complexity of human cognition, amongst its diverse cognitive functions, the metaphor enables the expression of that which is not literally expressible, such as internal states and emotions, or that of which one has not direct experience, relating it to more familiar or better known spheres (Cacciari, 1998, 1999; Glucksberg and Keysar, 1990; Ortony, 1980). Departing from the conviction that language is a complex and stratified faculty in which different cognitive abilities meet, my hypothesis is that the understanding of metaphoric language presupposes the interaction of a series of correlated components, such as the processes involved in linguistic elaboration, those responsible for perceptive analysis and imaginative mechanisms.

In my observations regarding the metaphor-imagination relation, I will use metaphors which are rooted in the perceptive-corporeal experience[1] as mediators between the physical dimension and the linguistic one, as a "cognitive bridge"

between experience and language. This type of analysis will enable me to show how imagination carries out the function of *trait d'union* between experience and the cognitive use of metaphor.

2. Principles of Cognitive Linguistics

A long linguistic-philosophical tradition, extending from structuralism to more recent generativist approaches, albeit with quite varying results, has studied semantics in accordance with an intra-linguistic and autonomist analysis of meaning. These concepts have no space for imagination, intended as a creative capacity or as a capacity to produce mental images or other structures of non-propositional formats. One of the factors ignored by these theories is the role played by the corporeal and perceptive dimensions in the construction of linguistic meaning.

In the Cognitive Linguistics (CL) perspective, which proposes to clarify cognitive processes at the base of the phenomena of meaning, it is not possible to talk about the autonomy of semantics. Language is bound by other aspects of human cognition: there is a close link between the plane of linguistic phenomena and that of non-linguistic faculties, that is the mechanisms of comprehension, of memory, of categorization, perceptive abilities and the cognitive capacity at the base of interaction between human beings and the environment on the one hand and with other human beings on the other. From this viewpoint, the mind is not a pre-established structure independent of human experience, rather it is "embodied" (Lakoff, 1987): a relevant, even if not total, part of our knowledge and cognition is based on the body and on its specific perceptive, motor and kinesthetic characteristics. From this perspective, since linguistic abilities are based on more general cognitive capacities, the metaphor is also studied in relation to the complex functioning of the human mind.

3. Embodiment and cognition

A close connection between language and embodiment is manifested, at a superficial level, in our daily linguistic production: expressions like "lately things have been going well/badly for me", show how, in occidental languages, movement is an important component in conceptualizing not typically motory experiences. Just as it is evident in the passage from the physical sense of grasping to the figurative sense of understanding[2] ("I didn't grasp that concept"); from the field of vision to that of cognition ("I don't see the reason in it", "a point of view") and how it is necessary to represent time in spatial terms: we say "in, for, within two hours", as we would say "in, for, within two kilometres"; "before two o' clock, after two o' clock, until two o'clock" as we would say "before the traffic lights, after the traffic lights, until (you reach) the traffic lights" (Cardona, 2001).

Studies relative to languages, which are not just occidental (Givon, 1979; Clark, 1978; Claudi and Heine, 1986; Lyons, 1967) show that the formation of the grammatical system of a language depends on cognitive and linguistic processes which make reference to constructions and words endowed with a specific

perceptive meaning. In synthesis, body, movement, perception, and spatial references constitute, in all cultures, fundamental reference schemes for conceptualizing different aspects of reality.

However, the relationship between language and embodiment is not just superficial. The perceptive, motor, spatial, and body components structure reality pre-linguistically, they constitute the "primitives" of the human conceptual system and are at the base of the configuration of sense (Gibbs and Colston, 1995; Johnson, 1987; Lakoff, 1987; Mandler, 1998, 2004). As early as the intrauterine period the baby is moving and has sensory and somatosensory experience of the consequences of his own movements. In the weeks following birth, through his own movements, he defines his own portion of space around his body, building a motor, and then a visual representation of space. The human brain, beginning with the initial phases of development, actively uses these basic motor, perceptive and spatial representations which constitute a sort of boundary to which language is subjected (Benelli *et al*., 1980; Fogassi, 2005; Mandler, 1993).

From this theoretic background, the merit of CL is having shown not only in which way some characteristics of categorization and of perception are reflected in language at a lexical and grammatical level (Casonato and Cervi, 1998), but also the systematic way in which concrete expression and spatialization processes come into play in figurative language and thought. This reasoning has contributed in by-passing the dualism of mind and body once and for all: the turning point is in having indicated an area which joins the motor-perceptive level, the imagination, memory and the more refined production of verbal language. This is the area of the metaphor.

4. Metaphor in CL

It is in the categorization and in the formation of meaning processes which, according to CL, metaphors which are not casual analogies, but which reflect the systematic link between bodily experience and human cognition come into play. From this point of view the metaphor is a cognitive device which enables the understanding of an abstract conceptual domain thanks to making reference to a concrete conceptual domain (Johnson, 1987; Lakoff, 1993; Lakoff and Johnson, 1980; Lakoff and Turner, 1989). The metaphor is therefore able to reduce to the concrete and the simple those domains which are difficult to represent cognitively for their abstraction and contingent lack of information, allowing the individualization of the physical-perceptive components of cognition.

Let us think of metaphorical extension in spatial terms and concepts to describe time ("an endless night", "the days to come"). To transform the notion of time into an operative schema necessitates an analogical procedure which will anchor it to something of which one has direct experience: space (Cardona, 2001).

It is therefore thanks to the metaphor, based on similarity and relationships between objects, events and experience, that we give "body" to abstract concepts, extending, through the mechanism of projection (*mapping*)[3], what is known to what is unknown by constructing a mental model of a reality which would be otherwise indescribable, in that it is not immediately perceived by the

senses. From this point of view the metaphor is intended as a projection on a linguistic and conceptual level of embodied configurations of varied nature – perceptive, motor, spatial – which Johnson (1987) and Lakoff (1987) call "image schemata".

5. From image schemata to metaphors

Reflecting on the notion of "imagination", it emerges that it is intuitively considered as an essentially creative, inventive activity, often used as a synonym for fantasy. In the cognitive context, and above all in relation to the metaphor, this dimension is neglected to give space to that which one can consider the subcomponents of imagination, that is mental images and image schemata. Cognitive sciences have given great importance to the notion of imagery, as a cognitive capacity to produce mental images (Ferretti, 1998). The scientific community does not agree on the relationship between metaphor and imagination, as intended in this meaning. Some researchers (Harris, Lahey and Marsalek, 1980; Verbrugge, 1977) link these two phenomena, by explaining the use of mental images in the metaphors coding and by assuming a close relationship between metaphorical and perceptive representation. From the experiments carried out by other researchers (Riechmann and Coste, 1980; Moran, 1989), the fact that the comprehension of metaphorical phrases can take place without the formation of perceptive representations such as mental images emerges. The thesis according to which the semantic domains in play in a metaphor are merged into a mental image can be intuitively convincing when applied to a metaphor such as "My doctor is an elephant" (Riechmann and Coste 1980: 183), in which the domains "people" and "elephants" are linked by an image which incorporates both.

The situation changes if we consider metaphors which are not simply predicative (such as "A is B"), but are more complex (such as "I'm on top", "Stock exchange shares have gone down/gone up", "He's come a long way in life"). They are equally comprehensible but are not represented in mental images in an immediately perceptive way. According to Moran (1989) the understanding of the meaning of a word does not coincide with having a mental image. A metaphor can definitely give rise to a play of images in the mind of the listener, but there is no reason to say that it is always so, and that this leads to a total understanding of the metaphor. Analysing the relational and complex metaphor, one realizes that the notion of analogy, central to the metaphorical mechanism, cannot be reduced to a mere morphological similarity, rather the experiential dimension of the domain in play in the metaphor must be called into question.

I therefore believe that one does not get to understand how a metaphor – consisting of a dynamic transfer process – works by studying mental images that are too specific, fixed and rigid. One therefore searches for the links between metaphor and imagination in a more indirect way. CL helps us in this task by affirming that metaphors are based on not logical-propositional but imaginative cognitive components: "the image schemata". An image schema is defined as a "dynamic pattern that functions somewhat like the abstract structure of an image, and thereby connects up a vast range of different experiences that manifest the same recurrent structure" (Johnson 1987: 2).

The schemata in themselves do not correspond to a precise mental image, but maintain the topological structure of the entire perceptive experience. They are recurring models of embodied experience, organizational centres of knowledge which structure an important amount of information into non-propositional form. This information emerges from sensory-motor activities such as manipulating objects, orientating oneself spatially, directing one's perceptive attention at various goals. Human beings organize and structure their relationships with the external environment through image schemata in a significant way before using conceptual or linguistic structures, thanks to the fact that the physical-perceptive experience is not chaotic, but is intrinsically organized even before concepts come into play. In this sense the way in which the body interacts with the environment is connected to the way in which we conceptualize reality.

Some examples of such schemata are UP-DOWN, FRONT-BACK, linked to the spatial orientation, CONTAINER, originating with the fact of experiencing our body as a container, PART-WHOLE, SOURCE-PATH-GOAL. Projected metaphorically from concrete conceptual domains, the image schemata enable the conceptualization of abstract domains. The metaphor, therefore, placing itself on an intermediate level between concrete concepts and abstract ones, which do not directly emerge from embodied experience, thanks to the image schemata enables the elaboration of the first for the successive structuring of the second.

Let us dwell, for the purpose of giving examples, on the image schema UP-DOWN derived from our perceptive experiences connected with the erect posture of *Homo Sapiens*, and on the metaphors related to it. Standing erect determines the coordinates of one's orientation in a universal way. To live in a world without gravity and not to be endowed with an erect stance would lead us to perceive space and the environment around us in a different way, even changing the way in which we conceptualize and talk about space. According to the hypothesis which I am analysing, through the metaphor we can "work" on schemata so as to construct models of correspondence between different conceptual domains. For example, in occidental languages, the image schema UP-DOWN, connected to the physical sense of vertical orientation, is projected metaphorically on many abstract domains, such as happiness and health in such a way that all which is correlated on physical and psychic well-being is conceptualized in terms of verticality and is understood to go in an upwards direction, while bad psychological and physical conditions go in the opposite direction ("I'm on top", "I'm depressed" are based on the orientation metaphor[4] HAPPINESS IS UP, SADNESS IS DOWN)[5]. This structuring depends on the fact that the schemata are organized internally in a precise way: each schema is not only made up of a number of parts and of relations between these parts (Johnson, 1987) but is also highly flexible in the sense that it is applicable to different conceptual domains. In virtue of its great flexibility and adaptability to different conceptual spheres, the image schemata can be considered to be the true «"embodied" anchors of the entire conceptual system» (Hampe 2005: 2) capable of taking the structuring process of abstract thought into account.

In the end, therefore, imagination is none other than the capacity to activate and coordinate schemata from which the same linguistic activity springs, through

metaphorical projection. In this sense imagination carries out the function of *trait d'union* between experience and cognitive use of the metaphor.

6. Cognitive plausibility of image schemata

The image schemata represent a sort of adhesive between embodied experience, thought and language and are therefore indicative of the existing continuity between our physical interaction in the world and our activities of imagination and thought. A series of empirical studies show that image schemata are not mere abstractions constructed for theoretical goals but are psychologically real, in the sense that they carry out a cognitive role in the different comprehension processes (Gibbs and Colston, 1995).

Gibbs (2005) claims that despite the dynamic nature attributed to image schemata in Johnson's classical elaboration (1987), it has been traditionally talked about in a static and disembodied manner. Instead, he claims that reasoning based on image schemata does not consist in the activation of already existing representational entities, but in the embodied simulation of events: "Image-schematic reasoning does not simply mean doing something with one's mind, but constructing a simulation of experience using one's body". (Gibbs 2005: 115) From this viewpoint the image schemata are simulators, "imitators" of real actions – in the sense that the schemata are processed continually during our effective perceptive and cognitive experience – and of potential actions which a person can carry out:

> image schemas are as much created in the moment, even in the absence of bodily movement, as they are retrieved from long-term memory. This perspective on image-schematic reasoning suggests that many aspects of perception, cognition, and language use are intimately tied to both real and imagined bodily action (Gibbs 2005: 117).

In other words, "understanding" is "imagining": the same linguistic comprehension, which emerges from image schemata, is based on simulation processes which reproduce embodied models of what we want to mean. As we will see better later, even though actions are not effectively carried out, neural circuits connected to image schemata are activated in many acts of linguistic comprehension "as if" – and it is in this sense that one talks about simulation – such actions were really carried out.

Mandler (1998, 2004, 2005) also supports the cognitive plausibility of image schemata: preverbal concepts – at the foundation of human thought and language – are well represented in image schemata which emerge with early infancy. He assumes a cognitive device which he calls "perceptual meaning analysis" which has the function of extracting the spatial and motor structure of events in an image-schematic form which enable to represent them. Let us think of the case of movement. As early as in their first months of life babies grasp the difference between an object which begins to move without anything coming into contact with it ("self motion") and an object which will move only when something comes into contact with it ("caused motion"). Furthermore, babies are able to distinguish between animated objects which move and interact with other objects

both directly and from a certain distance, and inanimate objects which move only when they are moved and are incapable of interacting with other objects at a distance. These discriminations, taken in isolation, are simple perceptive data which must be reorganized and structured in a new format. This is the task of "perceptual meaning analysis", a procedure which produces a simplified representation of the information provided by the perceptive system. The products of this operation have a very precise format, that of the image schemata.

In the case which I am considering, that of movement, the image schemata which emerge from the "perceptual meaning analysis" are those of ANIMATE MOTION, INANIMATE MOTION, SELF MOTION, CAUSED MOTION (Mandler 1993, 2004, 2005). While growing, children give their attention to other aspects, such as spatial relations, which enable the development of further categorizations and ever more complex versions of schemata which will play a decisive role at a linguistic level:

> For example, the SOURCE-PATH–GOAL image schema shows a developmental timeline of increasing complexity throughout the first year. At five months infants are able to attend to the goal of the path traced out by a human hand reaching toward an object (Woodword, 1998); then, at nine months they can distinguish between a hand grasping an object and a hand resting upon it (Woodward, 1999); while at twelve months infants are able to selectively attend to objects by following changes in the direction that a caregiver points or looks (Woodward and Guajardo, 2002) (Rohrer 2005: 175).

The evidence to support the role played by image schemata in cognitive processes does not just come from cognitive psychology spheres but also from neuroscience which shows how they are neurobiologically founded.

7. Neurobiological plausibility of image schemata

The theoretic model which I have presented has received interesting confirmation in experimental areas of cognitive neurosciences which propose models of structure and neural elaboration which support the thesis of the embodied mind.

Several researches (Coslett, Saffran and Schwobel, 2002; Hauk, Johnsrude and Pulvermüller, 2004; Schwoebel and Coslett, 2005) show how areas of the sensorimotor cortex are crucial in the semantic comprehension of expressions and terms connected with *embodiment*. According to this hypothesis, semantic comprehension comes about across structures, such as image schemata, which are localized in the same cortical area noted in sensorimotor activity. According to Rohrer (2005), the image schemata are dynamic activation patterns that are shared across neural maps of the sensorimotor cortex; he has demonstrated the neurobiological plausibility of image schemata and their role in linguistic comprehension asserting that the image-schematic functions of the sensorimotor cortex structure linguistic and even metaphorical expressions. Applying fMRI ("functional magnetic resonance imaging") and ERPs ("event related potentials") has shown that both the literal language and metaphorical one, relative to the body and to embodied actions, involve some areas within the primary sensorimo-

tor, pre-motor and the more secondary somatosensory cortices, which are not simply activated but are functionally involved in the semantic processing of those expressions.

Rohrer's proposal (2001) is that during the metaphorical processing the sensorimotor areas corresponding to sensorial dimensions in play in the metaphor are activated: for example, the visual metaphors activate visual areas, the spatial metaphors the spatial ones, the motor or tactile ones their corresponding areas, and so on. Therefore, when we read about grasping an object (even in the metaphorical sense), in a certain sense we mentally *simulate* and imagine carrying out that action, activating the same cortical areas.

The idea that some processes of comprehension are based on simulation mechanisms, which I have mentioned also with regard to Gibbs's researches (2005), leads us to studies carried out on mirror neurons (Rizzolati and Sinigaglia, 2006): from a neurological point of view, observing actions carried out by others activates cortical circuits responsible for the execution of those actions. Even if the actions observed are not effectively reproduced, the motor system is activated "as if" such actions were really carried out. In other words the observation of an action involves its internal simulation. According to experiments by Tettamanti *et al*. (2005), this system plays an important role in the comprehension of sentences connected to actions: listening to sentences which describe actions carried out by hands, legs and mouth involves the activation of visual-motor circuits which are used for the observation and execution of those actions.

In order to test the connection between metaphor and embodiment, Gallese and Lakoff (2005) propose to study the activation of the sensorimotor cortex testing if, for example, in the case of the concept of "grasping", the activation of parietal pre-motor circuits comes about not only effectively when one carries out the action of grasping an object, but also when one understands metaphorical sentences containing the concept of grasping. The two researchers suppose a positive response to this test on the basis of studies by Tettamanti *et al*. (2005). The idea is that the same neural substratum used in imagining and carrying out actions is utilized in comprehending sentences, even metaphorical ones, which describe those actions: "both imagination and (language) understanding are based to some extent on the "mental simulation" of sensory-motor experienced, the outlines of which are captured by image schemas" (Hampe 2005: 7).

This type of research suggests that to discover the cerebral structures utilized when we speak about an event, we must investigate the cerebral structures used when we observe and physically take part in that event. Different studies show that in the brain the schemata are localized in secondary sensorimotor areas rather than the primary ones (Gallese and Lakoff, 2005). While primary sensorimotor areas are connected to information processing relative to particular modality (visual, auditory, tactile, motor) and are directly connected to receptors and effectors linked to that particular modality (eyes, ears, skin, muscles), the secondary sensorimotor areas are connected to primary areas. In particular, some neurons in these secondary areas are sensitive to more than one modality of information, that is they are inter-modal: for example, the secondary motor areas integrate motor, visual and somatosensorial modalities in order to carry out motor actions.

The same image schemata have a cross-modal nature. This is demonstrated by the research carried out by Gallese and Lakoff (2005) and Dodge and Lakoff (2005), according to which the schemata are computed by neural circuits used in multi-modal sensorimotor operations. In fact, the experiments show that a variety of neural circuits is involved in the computation of the image schemata. As the brain is highly interconnected, the neural circuits which compute an image schema do not operate in an independent manner but are connected to other cerebral areas, including the neural circuits which compute other image schemata. Thus we can find the same image schema at the base of experience and different metaphorical formulations:

> the presence of image-schematic structure in a given area of the brain does not necessarily mean that this is the only area that computes this image schema. In fact, it seems reasonable to suppose that the more generally applicable a schema is, the more areas it is likely to "appear" in (Dodge and Lakoff 2005: 76).

It's clear therefore that in the human mind, which is embodied, the perceptive conceptual and linguistic structures share some processes: mechanisms responsible for perception, motion, spatial orientation, manipulation of objects, are also involved in linguistic processes, including metaphorical ones (Lakoff and Johnson, 1999).

The hypothesis according to which the comprehension of metaphorical language presupposes the interaction of a series of correlated components is supported by the studies in the field of neurosciences, which show that linguistic elaboration is not the exclusive domain of the left cerebral hemisphere, as claimed by a long tradition of research, but also the right hemisphere, specialized in imaginative-spatial functions, carries out very precise tasks in the comprehension of language, even when it is metaphorical. The dominant conception today rejects the thesis of the spatial localization of cognitive functions: both hemispheres contribute to the "construction" of the global meaning of a linguistic message and therefore the processing of metaphorical expressions involves vast neural systems which are distributed bilaterally. According to the studies of Bambini (in press), with respect to literal sentences, metaphors mainly activate the areas of Broca and Wernicke (left inferior frontal gyrus and left middle temporal gyrus), noted for syntactic and semantic processes; the homologous right of the areas of Broca and Wernicke and insula bilaterally, recently brought to light in various types of linguistic tasks; the left superior temporal sulcus, precuneus bilaterally, and other areas to be referred to cognitive processes which are not purely linguistic, such as theory of mind, visual imagination and episodic memory.

The metaphorical comprehension is therefore a complex cognitive phenomenon based on the interaction between linguistic processes and other cognitive processes, such as imaginative ones based on image schemata.

8. Synesthesia and cross-modality

When analysing daily linguistic production, the pervasiveness and productivity of some metaphorical sources such as the body and spatial, motor and perceptive components emerge. I believe that special attention should be directed at metaphors linked to these components because they are fundamental in understanding how the human mind creates links of a superior order between different cognitive and sensorial domains, enabling the discovery of further similarities and bringing the existence of hidden relationships to light. We could think of the metaphor as a "bridge" between language and corporeity (Beck, 1978, 1987; Cacciari, 2005), an operative capable of creating cognitive and neural links and of carrying out "leaps" between different cognitive domains, enabling movement from abstract to concrete in order to attribute a perceptive-embodied dimension to abstract domains. Since motor-perceptive-physical experience is the basis from which concepts are formed, the metaphor interests us as a "cognitive bridge" (Beck, 1978, 1987; Cacciari, 2001) capable of introducing a "sensorial logic" to a conceptual level (Beck, 1978). In fact the very perceptive-imaginative aspects determine the specificity of metaphors.

This notion of the "bridge" has an interesting key of interpretation in the synesthetic metaphor, which consists in nominating a perceptive experience, typical of a sensory organ, utilizing terms which refer to a different sensorial system (for example, "piercing noise", "acute pain", "bitter decision"). I believe, following the studies by Ramachandran (2003) and Ramachandran and Hubbard (2001a, 2001b), that synesthesia is a crucial phenomenon in the comprehension of the metaphor, for the manner in which the brain elaborates sensorial information and uses it to create abstract links between inputs which are apparently not correlated. Understanding the functioning of the synesthesia would let us throw some light on the tendency of the human mind to produce metaphors. The synesthetic metaphor, as a mediator in the rapport between perception and language, is in fact paradigmatic of the tendency of sensory experience to depend on its diverse modalities in a transversal way (Cacciari, 2005). In other words, the synesthetic metaphor linguistically reflects some cross-modal transfers of typical properties of synesthetic perception. This capacity for interconnection, as we have seen, is characteristic of the human cerebral structure itself and is at the base of the functioning of the neural circuits connected to image schemata.

I can therefore hypothesize that the transferral process, which characterizes every sort of metaphor, uses the cross-modality typical of experience and the interconnectivity typical of the functioning of the brain. Just as through synaethesia links are created between different senses, the metaphor creates formerly inexistent relations amongst concepts making different phenomena converge. In this sense I can suppose that not only do many metaphors have perceptive-embodied roots, but that the capacity to create metaphors is based on multimodal characters of image schemata and on the synesthetic character of perception. The metaphor is therefore based on a profound perceptive, cognitive and neural interconnectivity.

Conclusion

In this essay I have considered image schemata emblematic of the way in which the meaning, imagination, reasoning, and thought emerge from the embodied interaction with the environment: "[s]tructures of *perceiving* and *doing* must be appropriated to shape our acts of *understanding* and *knowing*". (Johnson 2005: 16). In this theoretic context, the metaphor seems adequate to me to act as a bridge between language, domains which are difficult to express and the perceptive sphere, thanks to its capacity to base itself, to extend expressive and cognitive capacities of language, in the perceptive-embodied world and to anchor itself on a great plasticity and perceptive and cognitive interconnectivity, which needs further investigation (Cacciari, 2005). In fact, I believe that future investigations of the imaginative basis of the metaphor must concentrate on a question of empirical and theoretic research, which has been neglected up to now, that is the analysis of relationships between synesthesia and image schemata. For example, this would deal with understanding if the cross-modality of synesthetic perception can motivate image schemata since these are computed by active neural circuits in multimodal sensorimotor operations. This type of analysis would throw some new light on metaphorical processes further clarifying in which way the mind, faced with the difficulties of verbally categorizing the experience – which is always cross-modal – turns to the metaphor, a precious device which can reveal the complexity of rapports between perception and language and between experience and thought, presenting itself as the manifestation of an intricate interwoven network of sensorial, cognitive and neural processes.

Endnotes

[1] Lakoff and Johnson (1980) admit that it is difficult to trace a clear boundary line between the physical and cultural bases of a metaphor. In fact each experience occurs within a vast background of cultural presuppositions and metaphors are characterized by an interlacing of physical-perceptive and cultural data. However, if we give our attention to the conceptual domains in which metaphors are constructed, it would seem legitimate to consider, for example, that the domain of show business is more cultural and that of spatial orientation more physical-perceptive. I will be dwelling upon this latter sort of domain.

[2] The word "comprendere" ("to understand") in Italian reveals its motor and tactile origins in the etymology "com-prehendere" ("to take", "to grasp"). Just as the Greek word "katalambano" means "I take" and "I understand", and the Latin word "intelligere," made up of "inter" and "legere" ("to collect", "to choose"), means "to understand". It is as if the act of taking was equipped to place itself in front of the object to observe and manipulate it in order to collect some of its characteristics.

[3] For an explanation of the mechanism of *mapping*, see Lakoff, 1993.

[4] For an analysis of orientation, ontological and structural metaphors, cf. Lakoff and Johnson 1980.

[5] It's clear that the different cultures organize their own conceptual systems according to different goals and that they establish on which perceptive, spatial and embodied aspects abstract concepts will be structured. That which is common to all cultures is the cognitive strategy utilized to understand abstract concepts in terms of the concrete ones. The fact that all cultures attribute centrality to the embodied metaphors (Douglas, 1973; Needham, 1973; Onians, 1954) is not therefore just chance but indicates precise processes of thought.

References

Bambini, V. (in press). *Le metodiche delle neuroscienze nello studio dei fenomeni pragmalinguistici. Dati fMRI ed ERPs sulla comprensione della metafora*. Atti del XL Convegno annuale della Società di Linguistica Italiana, Vercelli 21-23 settembre. Roma, Bulzoni.

Beck, B. (1978). The metaphor as a mediator between semantic and analogic modes of thought. *Current Anthropology*, 19, 83-97.

Beck, B. (1987). "Metaphor, cognition and Artificial Intelligence". In R.S. Haskell (Ed.) *Cognition and symbolic structure: The psychology of metaphoric transformation*, Norwood, NJ, Ablex Publishing, pp. 9-30.

Benelli, B., D'Odorico, L., Levorato, M.C., Simion, F. (1980). *Forme di conoscenza prelinguistica e linguistica*. Firenze, Giunti Barbera.

Cacciari, C. (1998). "Why Do We Speak Metaphorically? Reflections on the Functions of Metaphor in Discourse and Reasoning". In A. Katz *et al*. (Eds.) *Figurative language and thought*, Oxford, Oxford University Press, pp. 119-157.

Cacciari, C. (1999). La metafora: un ponte fra il linguaggio e l'esperienza percettiva. *Lingua e Stile*, XXXIV, 2, 159-200.

Cacciari, C. (2001). *Psicologia del linguaggio*. Bologna, Il Mulino.

Cacciari, C. (2005). "Il rapporto tra percezione e linguaggio attraverso la metafora". In A.M. Lorusso (a cura di) *Metafora e Conoscenza*. Milano, Bompiani, pp. 321-348.

Cardona, R.C. (2001). *I sei lati del mondo. Linguaggio ed esperienza* (Terza ed.). Roma-Bari, Editori Laterza.

Casonato, M., Cervi, M. (a cura di) (1998). *Elementi di Linguistica Cognitiva*. Urbino, Edizioni QuattroVenti.

Clark, E. (1978). "Locationals: Existential, Locative and Possessive Constructions". In J. Greenberg (Ed.) *Universal of Human Language*, vol. 4. Stanford, Stanford University Press, pp. 85-126.

Claudi, U., Heine, B. (1986). On the Metaphorical Base of Grammar. *Studies in Language*, 10, 297-335.

Coslett, H.B., Saffran, E.M., Schwoebel, J. (2002). Knowledge of the human body: a distinct semantic domain. *Neurology*, 59, 357-363.

D'Arcy Thompson, W. (1961). *On Growth and Form. An abridged Edition edited by J.T. Bonner*. Cambridge, Cambridge University Press; (trad. it. *Crescita e forma*, Edizione ridotta a cura di John Tyler Bonner, Torino, Boringhieri, 2001).

Dodge E., Lakoff G. (2005). "Image schemas: From linguistic analysis to neural grounding". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 57-91.

Douglas, M. (1973). *Natural symbols*. New York, Vintage Books.

Ferretti, F. (1998). *Pensare vedendo. Le immagini mentali nella scienza cognitiva*. Roma, Carocci.

Fogassi, L. (2005). "La metafora nel cervello: il punto di vista neurofisiologico". In A.M. Lorusso (a cura di) *Metafora e conoscenza*. Milano, Bompiani, pp. 349-361.

Gallese V., Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22, 3-4, 455-479.

Gibbs, R.W.Jr. (1994). *The poetics of mind, Figurative thought, language and understanding*. Cambridge, Cambridge University Press.

Gibbs, R.W.Jr. (2005). "The psychological status of image schemas". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 113-135.

Gibbs, R.W.Jr., Colston, H.L. (1995). The cognitive psychological reality of image schemas and their transformations. *Cognitive Linguistics*, 6, 347-378.

Givón, T. (1979). *On Understanding Grammar*. New York, Academic Press.

Glucksberg, S., Keysar, B. (1990). Understanding metaphorical comparison: Beyond similarity. *Psychological Review*, 97 (1), 3-18.

Greenberg, J. (Ed.) (1978). *Universal of Human Language*. Vol. 4. Stanford, Stanford University Press.

Hampe, B. (Ed.) (2005). *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter.

Harris, R.J., Lahey, M.A., Marsalek, F. (1980). "Metaphors and Images: Rating, Reporting, and Remembering". In R.P. Honeck, R.R. Hoffman (Eds.) *Cognition and Figurative Language*, Hillsdale, NJ, Erlbaum, pp. 163-181.

Haskell, R.S. (Ed.) (1987). *Cognition and symbolic structure: The psychology of metaphoric transformation*. Norwood, NJ, Ablex Publishing.

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

Honeck, R.P., Hoffman, R.R. (Eds.) (1980). *Cognition and Figurative Language*. Hillsdale, NJ, Erlbaum.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago, The University of Chicago Press.

Johnson, M. (2005), "The philosophical significance of image schemas". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 15-33.

Johnson, M.G., Malgady, R.G. (1980). "Toward a Perceptual Theory of Metaphoric Comprehension". In R.P. Honeck, R.R. Hoffman (Eds.) *Cognition and Figurative Language*. Hillsdale, NJ, Erlbaum, pp. 259-282.

Katz, A., Cacciari, C., Gibbs, R.W.Jr., Turner, M. (1998). *Figurative language and thought*. Oxford, Oxford University Press.

Kuhn, D., Siegler, R.S. (Eds.) (1998). *Cognition, Perception, and Language*. New York, Wiley.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago, The University of Chicago Press.

Lakoff, G. (1993). "The Contemporary Theory of Metaphor". In A. Ortony (Ed.) *Metaphor and Thought*. Cambridge, Cambridge University Press, pp. 202-251.

Lakoff, G., Johnson, M. (1980). *Metaphors We Live by*. Chicago, The University of Chicago Press; (trad. it. *Metafora e vita quotidiana*, Milano, Bompiani, 1998).

Lakoff, G., Johnson, M. (1999). *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. New York, Basic Books.

Lakoff, G., Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago, University of Chicago Press.

Langacker, R. (1976). Semantic Representations and the linguistic relativity hypothesis. *Foundations of Language*, 14, 307-357.

Langacker, R. (1987). *Foundations of Cognitive Grammar*. Vol 1. Stanford, Stanford University Press.

Lorusso, A.M. (a cura di) (2005). *Metafora e conoscenza*. Milano, Bompiani.

Lyons, J. (1967). A Note on Possessive, Existential and Locative Sentences. *Foundations of Language*, 3, 390-396.

Mandler, J. (1993). On concepts. *Cognitive Development*, 8, 141-149.

Mandler, J. (1998). "Representation". In D. Kuhn, R.S. Siegler (Eds.) *Cognition, Perception, and Language*, New York, Wiley, pp. 255-308.

Mandler, J. (2004). *The Foundations of Mind: Origins of Conceptual Thought*. Oxford, Oxford University Press.

Mandler, J. (2005). "How to build a baby: III. Image schemas and the transition to verbal thought". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 137-163.

McMahon, T.A., Bonner, J.T. (1983). *On size and life*. New York, W.H. Freeman and Company; (trad. it. *Dimensioni e vita*, Bologna, Zanichelli, 1990).

Mazzeo, M. (2005). *Storia naturale della sinestesia. Dalla questione Molyneux a Jakobson*. Macerata, Quodlibet Studio.

Moran, R. (1989). Seeing and Believing: Metaphor, Image, and Force. *Critical Inquiry*, 16, 87-112.

Needham, R. (1973). *Right and left: Essay on dual symbolic classification*. Chicago, University of Chicago Press.

Onians, R.B. (1954). *The origins of European thought about the body, the mind, the soul, the world, time, and fate*. Cambridge, Cambridge University Press.

Ortony, A. (1980). "Some psycholinguistic aspects of metaphor". In R.P. Honeck, R.R. Hoffman (Eds.) *Cognition and figurative language*. Hillsdale, NJ, Erlbaum, pp. 69-83.

Ortony, A. (Ed.) (1993). *Metaphor and Thought*. Cambridge, Cambridge University Press.

Popova Y. (2005). "Image schemas and verbal synaesthesia". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 395-419.

Ramachandran, V.S. (2003). *The emerging mind*. United Kingdom, Profile Books Ltd; (trad. it. *Che cosa sappiamo della mente*, Milano, Mondadori, 2004).

Ramachandran, V.S., Hubbard, E.M. (2001a). Psychophysical investigation into the neural basis of synesthesia. *Proceedings of the Royal Society of London,* B, 268, 978-983.

Ramachandran, V.S., Hubbard, E.M. (2001b). Synesthesia, a window into perception, thought and language. *Journal of Consciousness Studies*, 8 (12), 3-34.

Riechmann, P.F., Coste, E.L. (1980). "Mental Imagery and the Comprehension of Figurative Language: Is There a Relationship?". In R.P. Honeck, R.R. Hoffman (Eds.) *Cognition and Figurative Language*. Hillsdale, NJ, Erlbaum, pp. 183-200.

Rizzolatti, G., Sinigaglia, C. (2006). *So quel che fai. Il cervello che agisce e i neuroni specchio*. Milano, Raffaello Cortina Editore.

Rohrer, T. (2001). *Understanding through the Body: fMRI and ERP studies of metaphoric and literal language*. Paper presented at the 7th International Cognitive Linguistics Association Conference, University of California at Santa Barbara, Santa Barbara, CA, July 2001.

Rohrer, T. (2005). "Image Schemata in the Brain". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 165-196.

Schwoebel, J., Coslett, H.B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience*, 4, 543-553.

Shaw, R., Bransford, J. (Eds.) (1977), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ, Erlbaum.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Vol. 1. Cambridge Mass., MIT Press.

Talmy, L. (2005). "The fundamental system of spatial schemas in language". In B. Hampe (Ed.) *From Perception to Meaning: Image Schemata in Cognitive Linguistics*. Berlin-New York, Mouton de Gruyter, pp. 199-234.

Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., *et al*. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273-281.

Tobias, P.V. (1982). *Man: the Tottering Biped*: *The Evolution of His Posture, Poise and Skill*. Committee in Postgraduate Medical Education. Univer-

sity of New South Wales (trad. it. *Il bipede barcollante. Corpo, cervello, linguaggio*. Torino, Einaudi, 1992).

Verbrugge, R.R. (1977). Resemblances in language and perception. In R. Shaw, J. Bransford (Eds.) *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ, Erlbaum.

Went, F.W. (1968). The size of man. *American Scientist*, 56, 400-418.

Erin Shaver - Brian Maniscalco - Hakwan Lau
Columbia University

AWARENESS AS CONFIDENCE

Abstract Lesion to the primary visual area in the brain abolishes visual awareness. And yet, patients with such lesions can perform forced-choice visual detection or discrimination tasks better than chance. In performing these tasks, however, they claim that they are only guessing, that is they have little confidence that they are correct. This paper considers whether this reported lack of confidence could help us to characterize the apparent lack of visual awareness. In other words, do confidence and awareness always go together? Or are there other consistent relationships between the two such that we could use confidence ratings as measures of visual awareness? We highlight why the approach of using confidence ratings to assess the level of visual awareness is better than the currently dominant approach, which is to use forced-choice performance itself as an index of awareness. But we also discuss potential pitfalls, both technical and conceptual, in interpreting confidence ratings in relation to visual awareness.

Key Terms Consciousness, Awareness, Confidence, Blindsight, Signal Detection Theory, Phenomenology.

1. Introduction

The history of using confidence ratings as a measure of awareness dates back to the 1880's, when Peirce and Jastrow were examining minute changes in pressure sensation. These authors asserted that a slight tactile pressure change could be detected consciously or unconsciously. The confidence rating measures not the detection of sensation but the subject's "feelings" about the sensation change. Peirce and Jastrow's experiment drew attention to the discordance between awareness and forced-choice responses, and they pointed to the fact that performance judgments may be correct at a rate above chance even when confidence ratings are at zero. In their conclusions, Peirce and Jastrow speculated that confidence might measure a "secondary sensation of a difference between the primary sensations compared". Forced choice judgments allow us to test those primary comparisons (change vs. no change), but a subjective report is necessary to ascertain the degree to which such judgments are consciously made (Peirce and Jastrow, 1885).

This is different from the current mainstream approach to studying visual awareness (Rees *et al*., 2002). In these studies, we typically require subjects to make forced-choice judgments on some aspects of the visual presentation (whether the target is there or not, or whether stimulus A or B is presented). Researchers typically take the accuracy of such forced-choice judgments as an index of visual awareness. If subjects can perform such tasks, we assume that they are visually aware of the stimuli. If they cannot, i.e. performance is at chance le-

level, we assume that they are visually unaware. The approach we are concerned with here is different, in that the experimenters collect subjective confidence ratings in addition to forced-choice judgments, and awareness is assessed based on these subjective ratings. This article evaluates whether this confidence-based approach is better than the current mainstream approach, and also highlights certain potential pitfalls in using it.

2. Two approaches to using confidence as a measure of awareness

To start, let us distinguish two ways of how one could use confidence ratings to measure awareness (Dienes, in press): (1) high confidence ratings (as determined by subjective report) can be equated with awareness directly; (2) a high correlation between confidence and performance can be used to determine level of awareness. The first approach is vulnerable to the subjects' overstating or understating their level of awareness, depending upon their personal criteria for confidence. Whereas one subject might report a high level of confidence based on a weak hunch, another might report low confidence due to discomfort with even a minor lack of certainty. The confidence rating taken alone is highly subjective and difficult to apply when comparing across subjects.

Proponents of the second approach attempt to address the above issue by measuring the correlation between confidence ratings and the performance of a forced-choice task. This approach focuses on whether subjects are able to place high subjective ratings at the right time, i.e. in trials where their judgments are correct, rather than whether subjects on the whole give high subjective ratings or not. It thereby somewhat alleviates the problem of uncontrolled overall biases that accompany using a liberal vs. conservative strategy. Essentially, the correlation approach focuses on a kind of metacognitive capacity, i.e. the ability to place the subjective ratings appropriately such that they reflect knowledge of when one is likely to be correct.

It is perhaps important to note here that by adopting either approach, one does not necessarily commit oneself to conceptually equating awareness and confidence, or otherwise reducing one in terms of the other. These should be taken as operational definitions, in the sense that if these conditions would capture the necessary and sufficient conditions for awareness, that would satisfy our purpose.

3. Motivations for using confidence as a measure of awareness

3.1. Forced-choice task performance measures are poor

It has previously been argued that performance in forced-choice tasks measures the capacity for visual processing, but not necessarily awareness (Lau, in press). Blindsight, for instance, is a classic example of a situation in which there appears to be a disassociation between performance ability and awareness (Weiskrantz, 1986, 1999). When forced to report on whether the stimulus (shown to the damaged visual field) is present or not, 80-90% of the time the subject will report correctly, indicating that the brain has processed the visual

information. The same subject, however, will report that he has no visual awareness in the damaged field, and that he does not perceive what he is seeing. In other words, we can have good forced-choice task performance and a complete lack of visual awareness at the same time, which means the former cannot be a good measure of the latter.

3.2. Confidence measures of awareness are not necessarily subjective

One apparent attraction of forced-choice task performance is that it can be objectively measured. Confidence ratings, on the other hand, seem to be entirely subjective. One could argue that this subjective nature may well be necessary; after all, consciousness is supposed to be subjective. However, the fact that confidence is subjective somewhat undermines its value as a good scientific measure.

In one recent experiment, subjects were asked to place monetary bets on their responses to forced-choice tasks in order to introduce an "objective measure of awareness" through post-decision wagering. (Persaud *et al*, 2007). After each trial, subjects were required to place a bet as to whether they are correct or incorrect in the forced-choice task. If they are confident that they are correct, presumably they would place a large bet, and if they lack confidence, the bet would presumably be small. The advantage of this approach is, unlike ordinary confidence ratings, now there is an objective answer as to how one should bet, i.e. how one could maximize the amount of money earned towards the end of the experiment. In this sense then, wagering can be used as a substitute for confidence ratings, thereby objectively measuring awareness.

It should be noted that although the approach of directly taking confidence ratings as an index of level of awareness might seem susceptible to subjective biases, the correlation approach as described in the last section essentially assesses a form of metacognitive capacity. One could argue that while the ratings are still subjectively made, the ability to place these ratings on the right trials is determined by objective cognitive factors. For this reason, some have even applied signal detection theory to refine this correlation approach (Kunimoto *et al*., 2001), to make it more objective.

Signal detection theory is a formalism that allows one to separate subjective decision criteria from cognitive capacity. For instance, in a visual detection task, individual differences in target detection rates could be explained by differences in overall willingness to give positive identification responses, in addition to differences in actual perceptual ability to detect a visual signal. Signal detection theory allows one to disentangle these two factors and assess the actual perceptual or cognitive performance without any confound arising from response bias. Kunimoto *et al*. (2001) applied the standard signal detection model to confidence judgments. They considered confidence judgments to be analogous to correct responses when they were congruent with actual performance (i.e. high confidence for correct responses and low confidence for incorrect responses). The idea was that this approach would allow one to assess a subject's ability to discriminate between his own correct and incorrect responses, independently of that subject's idiosyncratic tendency to give high confidence judgments. If successful, this approach could

capture the "subjective" aspects of confidence ratings in a measure more methodologically "objective" than taking subjective reports at face value.

4. Potential pitfalls

4.1. Wrongly applying standard signal detection models to the analysis of confidence ratings

Galvin *et al*. (2003) provided a theoretical framework for how to apply signal detection theory to confidence judgments about one's own performance. They referred to the primary task (e.g. to detect a visual stimulus) as the type 1 task, and the secondary task of judging one's confidence in one's type 1 response as the type 2 task. In their analysis they described what is the theoretically maximal type 2 performance, given a certain type 1 performance. Their analysis revealed that the type 2 task is theoretically constrained in ways different from the type 1 task.

For instance, signal detection theory posits that stimuli are evaluated along a decision axis, which quantitatively characterizes the perceptual or cognitive evidence or signal used to make judgments about the stimuli. In type 1 tasks, stimuli representations are modeled as being normally distributed along the decision axis. That is, members of a given stimulus class tend to occur most frequently at a particular location on the decision axis (the mean of the distribution), with some variance about this mean (the spread of the distribution) in such a way that the distribution is Gaussian. However, Galvin *et al*. (2003) showed that if type 2 tasks are performed using the same decision axis as type 1 tasks, then the type 2 "stimuli" (correct and incorrect responses) are not normally distributed along the decision axis. In fact, they are always systematically and radically different from standard normal distributions. Thus, any standard signal detection theory approaches that assume normal distribution of stimulus representations (such as the standard measure d') cannot be applied to type 2 tasks.

However, Kunimoto *et al*. (2001) assumed in their signal detection treatment of confidence judgments that correct and incorrect responses are distributed normally. If the analysis of Galvin et al. is correct, then Kunimoto et al.'s approach rests on theoretically dubious foundations. Indeed, Evans and Azzopardi (2007) discovered theoretical and empirical flaws in Kunimoto's application of d' analysis to confidence judgments. They showed that given a constant type 1 d', the type 2 d' as computed by Kunimoto et al. varies systematically with type 2 criterion. This theoretically predicted pattern was subsequently found to exist in actual human observers. In a visual target detection task, Evans and Azzopardi manipulated type 2 response criteria of their subjects in one experiment by changing the relative frequency of target stimuli, and in another by allowing subjects a predetermined number of responses at each level of confidence. Across both cases, Kunimoto et al's type 2 d' increased with increases in type 2 response criterion. Thus, it would seem that Kunimoto et al.'s analysis does not deliver the intended bias-free measure of type 2 performance. This is likely due at least in part to Galvin et al.'s observation that the assumptions signal detection theory makes with respect to type 1 tasks do not generalize to type 2 tasks.

Likewise, Galvin *et al.* (2003) showed that for an ideal observer, the range of values that type 2 performance could take is theoretically constrained by type 1 d', and further that the precise level of type 2 performance within that range that actually obtains is determined by the type 1 decision criterion. Because type 2 performance depends on the type 1 d' and decision criterion, it is misleading to straightforwardly evaluate type 2 performance. Rather, Galvin *et al.* (2003) endorse a more sophisticated approach wherein one evaluates type 2 performance by comparing it to the predicted type 2 performance of an ideal observer, given the type 1 d' and decision criterion of the actual observer.

All the technical specifics aside, perhaps the take home message is that signal detection theoretic analysis of confidence ratings can be complicated, with the theoretical details formally worked out only recently (Galvin et al. 2003). Some previous attempts have been shown to be seriously flawed. And research adopting such flawed analysis must be interpreted with caution.

4.2. Validity of the correlation approach

Even if we apply the correct formal model to assess the association between confidence ratings and performance, it is still questionable whether a high correlation between the two should mean a high level of awareness. In some sense, it is trivially true that if one's subjective ratings reflect different levels of effectiveness in information processing, one demonstrates some level of metacognitive awareness, i.e. one is aware of one's own performance. But what we are primarily concerned with here is sensory awareness, in particular whether a visual percept gives conscious phenomenology.

For instance, Kolb and Braun (1995) have studied different masking conditions under which a visual target is not phenomenologically registered, yet subjects could still locate the target at well above chance level. This result is interesting because this is essentially a demonstration of "blindsight in normal observers". The authors used a correlation approach to show that in these conditions, performance and confidence are uncorrelated. Later researchers (Robichaud and Stelmach, 2003 and Morgan *et al.*, 1997), however, have failed to replicate the result; they found instead that confidence and performance remained correlated across trials. This has led some authors to call this line of research utilizing visual masking a "fruitless line of inquiry". However, if the visual phenomenology associated with seeing the targets is actually absent, or at least nearly absent or much reduced compared to normal conditions (Morgan *et al.*, 1997), perhaps what we should challenge is the validity of the correlation approach, but not the masking phenomenon itself.

Indeed, it seems as if there are at least some cases where the correlation approach incorrectly predicts the presence or absence of awareness. According to the correlation approach, when confidence and performance are well correlated, the subject should be aware of the stimulus. However, in one recent experiment we have found that blindsight patient GY can have similar levels of correlation between confidence and performance in his good and blind fields, while still retaining the characteristically stark difference in visual awareness between the two fields (Persaud, Cowey, Maniscalco, Mobbs, Passingham and Lau, manu-

script in preparation). This demonstrates that a good correlation between confidence and performance is not sufficient for awareness.

Furthermore, it seems likely that the converse relationship need not hold either: a good correlation between confidence and performance is not necessary for awareness. Visual awareness can occur alongside a poor correlation between confidence and performance in instances where a subject's visual experience systematically fails to represent the external world faithfully in some respect. For instance, a subject administered a psychoactive, hallucinatory substance such as psilocybin while performing a motion detection task might visually experience false motion alongside actual motion. On this basis the subject might endorse incorrect responses with high confidence (mistaking hallucinatory motion for actual motion). The subject might also endorse correct responses with low confidence (not clearly perceiving the difference between actual and hallucinatory motion). Even cognitively normal subjects might fail to give confidence ratings in agreement with actual performance, e.g. if performing judgments on stimuli exploiting known visual illusions. In both scenarios, the failure to give appropriate confidence ratings arises from non-veridical visual awareness, rather than from a lack of awareness.

In sum, it seems as if awareness can be absent even when confidence and performance are well correlated, and likewise that awareness can be present even when confidence and performance are not well correlated. This calls into question the conceptual validity of the correlation approach, as such a correlation seems neither necessary nor sufficient for awareness.

4.3. Confidence may capture too much

The approach of taking confidence level directly as an index of awareness is not without its problems either. People in general vary in how much stock they put into their ability to act on a hunch or to feel their way through a situation. It is possible to imagine that some subjects may be quite used to "trusting their guts" even while not having vivid phenomenological basis for such intuitive feelings. One could argue that such gut feelings or intuitions are still part of the conscious experience. However, at the very least this highlights a problem for using confidence ratings as a measure for awareness in a particular perceptual modality, such as vision. High confidence in a visual task may not reflect high conscious visibility, because such confidence could be driven by non-visual experiences.

5. Closing remarks

The fact that conscious awareness is an essentially subjective phenomenon means that there is a need to take seriously subjective reports like confidence ratings. However, there are caveats in analyzing and interpreting such data. The possibility that one could use sophisticated analysis to make such measures more objective may seem appealing, but as we have highlighted, there are formal and conceptual problems associated with this approach. Interestingly, perhaps this indicates that one of the most basic ways of using subjective ratings, i.e. taking

the subjects' word for it as to whether they actually see (Weiskrantz, 1986), may well be the most straightforward and useful approach. If one can demonstrate a difference because this subjective report of visibility between two conditions where forced-choice performance is matched, one can be fairly confident that there is a real difference in terms of the level of awareness there (Lau and Passingham, 2006).

References

Dienes, Z. (in press). "Subjective Measures of Unconscious Knowledge", In R. Banerjee, B. Chakrabarti (Eds.) *Models of Brain and Mind: Physical, Computational and Psychological Approaches*. Elsevier.

Evans, S., Azzopardi, P. (2007). Evaluation of a 'bias-free' measure of awareness. *Spatial Vision*, 20 (1-2), 61-77.

Galvin, S.J., Podd, J.V., Drga, V., Whitmore, J. (2003). Type 2 tasks in the theory of signal delectability: discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10 (4), 843-876.

Kolb, C., Braun, J. (1995). Blindsight in normal observers. *Nature*, 377 (6547), 336-338.

Kunimoto, C., Miller, J., Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition*, 10, 294-340.

Lau, H.C., Passingham, R.E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *PNAS*, 103 (49), 18763-18768.

Lau, H.C. (in press). "Are we studying consciousness yet?". In M. Davies, L. Weiskrantz (Eds.) *Frontiers of Consciousness*.

Macmillan, N., Creelman, C. (1991). *Detection theory: a user's guide*. Cambridge [England]; New York, Cambridge University Press.

Morgan, M.J., Mason, A.J.S., Solomon, J.A. (1997). Blindsight in normal subjects? *Nature*, 385, 401-402.

Peirce, C.S., Jastrow, J. (1885). On small differences in sensation. *Memoirs of the National Academy of Sciences*, 3, 317-329.

Persaud, N., McLeod, P., Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, 10, 257-261.

Pessoa, L. (2007). Fear Perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision*, 7 (4), 10, 1-17

Rees, G., Kreiman, G., Koch, K. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3, 261-270.

Robichaud, L., Stelmach, L.B. (2003). Inducing blindsight in normal observers. *Psychonomic Bulletin & Review*, 10 (1), 206-209.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1-29.

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford, Oxford University Press.

Weiskrantz, L., Barbur, J.L., Sahraie, A. (1995). Parameters affecting con-

scious versus unconscious visual discrimination with damage to the visual cortex (V1). *PNAS*, 92, 6122-6126.

Wesikrantz, L. (1999). *Consciousness Lost and Found*. Oxford, Oxford University Press.

Gary J. Shipley
University of Kent

THE FICTIONAL AND THE REAL: THE DENNETTIAN SELF

Abstract Daniel C. Dennett claims that the self is nothing more than a fiction of the brain, an abstraction that has been promoted by evolutionary processes as a result of its biological and social beneficence. While concurring with Dennett with regard to *simple* selves, I argue for the existence of indeterminate and functional selves, and propose that such selves come about as a direct result of our believing in the reality of simple and thus fictional selves. In addition to this I shall contend that Dennett is confused as to the precise nature of the self, and as a consequence ends up fluctuating between two importantly different positions. My overall aim is to ameliorate Dennett's initial insights and hopefully come up with something more consistent and metaphysically palatable.

Key words Self, fiction, fictional construction, functional selves.

1. Introduction

The idea that the self is a fiction is hardly new[1]. However, what is novel about Daniel C. Dennett's account of the self is that instead of it being a grammatical or community fiction, it is a fiction of the brain, an abstract theoretical device promoted by evolutionary processes as a result of its biological and social beneficence. While I agree that there is no such thing as a simple self, I shall argue for the possibility of there being an indeterminate and functional self, and that such a self could come about as a direct result of our believing in the reality of a simple self – a self that turns out to be fictional. For there to be a real self there does not, *pace* Dennett, have to be some part of the world that directly corresponds to our (fictional) notion of the self. Dennett also proposes that the self is analogous to a fictional character, but it is not clear just how we are supposed to take this claim[2]. Is it that the self is fictional like the man Hamlet i.e., non-existent, or is it that the self is fictional like the role Hamlet i.e., existent and able to be occupied by existents while not being a factual account of the existence of some person who existed? It is my contention that Dennett in fact fluctuates between these two positions, so that it is ultimately unclear exactly what status the self holds for him. After attempting to shed some light on such confusions, I shall propose a reductionist and realist alternative to Dennett's eliminativist fictionalism that adapts the latter of these two notions of fiction and provides us with a clearer and more palatable account of the self.

2. The Self as a Fictional Construction

I will briefly explain Dennett's position before analysing certain aspects in greater depth. In *Consciousness Explained*, Dennett claims that the self is a theoretical fiction analogous to a centre of gravity. Both the self and the centre of gravity are claimed to be useful explanatory devices, while neither is identical to any part of the material world and so cannot be thought of as real. He asserts that humans create (or 'spin') selves in similar ways and for similar purposes (or non-conscious reasons) that spiders spin webs, beavers make dams and bower birds build bowers. A spider's web is not the result of conscious pursuit on behalf of the spider, but of the spider's "innate drives and proclivities" (Dennett 1993: 415). Similarly, the self of the human being is claimed to be a result of the human brain spinning a web of words and deeds around an innately posited organisational nexus. But – unlike the spider which simply exudes its web – the individual human being has a conscious role to play in the construction of its self, a role analogous to a beaver collecting materials for its dam or a bower bird collecting ornate objects for its bower. A human being labours over the manufacture of a successful self. It appears (although it is not altogether clear) as if the initial spinning results in what Dennett calls "a biological self", or that which reflects the initial realisation that there are reasons and interests that are 'mine', thus allowing the concern for self-protection to have an object with definable boundaries. Dennett talks of the self at two levels of abstraction: the "biological self" and the "psychological or narrative self". Neither of these selves is physically present in the brain, but both are, nevertheless, exceedingly efficient at attracting psychological and narrative assets. It is the biological self which is the primitive abstraction that enables the psychological or narrative self (itself another abstraction) to get a hold.

According to Dennett, this remarkable psychological self is constructed from "words" and "stories", for the self is nothing but a "web of discourses" (Dennett 1993: 416-18). You may well ask how it is that a self can be composed of words; after all, selves don't appear to be like sentences or paragraphs. However, in the sense that both can be written around a fictional focal point, there are ways in which they do. We do not, though, according to Dennett, consciously construct our own narratives, the process of construction being likened to the spider spinning his web, so that

> Our tales are spun, but for the most part we don't spin them; they spin us. Our human consciousness, and our narrative selfhood, is their product, not their source. (Dennett 1993: 418)

I think in order to get a proper grasp on what Dennett is getting at here it helps if we try to imagine the fictional self to be like a magnetic force that attracts particles to a single point; for it then becomes clear that its attracting stuffs to it need not involve any conscious endeavour. Recall that beavers, bower birds and humans differ from spiders, in that their creations have to be laboured over and the respective materials diligently collected. But then it follows from this that we do in fact play a role in the spinning of a self, despite not consciously constructing our narratives. This appears to involve Dennett in a contradiction.

However, this is not necessarily the case. For what Dennett means when he says that the forging of a psychological self is both something laboured over and something we have no conscious role in constructing, is simply that our labour is blind to its true (or evolutionary) purpose. We labour over the narrative that surrounds the psychological nexus, but no conscious effort is required in making that narrative one's own. Although the organisational nexus or self-representation around which stories are told is, according to Dennett, pure abstractum[3], it is nevertheless useful, because it

> […] gathers and organises the information on the topic of me in the same way other structures in my brain keep track of information on Boston, or Reagan, or ice cream (Dennett 1993: 429).

This self-representation is our centre of narrative gravity. But the stories and information that collect around Boston or Reagan will be, in all likelihood, a mixture of (primarily) factual knowledge and personal perspectives on the two items. No doubt the information collected is largely reliant on what happens in Boston's history and Reagan's life. So the stories we build around these points will (it would be fair to suppose) be largely reflective of the actual histories of Boston and Reagan. In which case, the self appears to be more analogous to autobiographical fiction than straight fiction – James Joyce's 'Stephen Dedalus' providing a more accurate analogy than Dennett's use of Melville's 'Ishmael'. The stories that surround the character Stephen Dedalus reflect actual occurrences in the life of James Joyce, whereas the stories that surround Ishmael are purely fictional, because they don't record what happened to any particular human being – at least, it is extremely unlikely that they do. But if the centre of narrative gravity is indeed analogous to the main character in autobiographical fiction, then the surrounding narrative or stories are about real things and real events, although the self to which the narrative is attributed, it being what the collection of tales are supposedly about, remains fictional. Our believing a purely representational self to be actual might well be what gives structure to the tales and imbues them with the ability to become a single narrative, but the structure for which it might be responsible is, nevertheless, real and a majority of the structured tales are themselves real – factual as opposed to fictional. So it could be claimed that a fictional self actually gives rise to what amounts to a factual and therefore straightforwardly real self.

Let me make my position more explicit by means of an analogy. Suppose that you tell a hundred people that there is a party in Hyde Park, when in reality there is no party. All hundred people are drawn to Hyde Park believing there to be a party. They all turn up with their bottles and food and as it turns out there is a party, just not the one they presumed was already in full swing. Each of the guests went to Hyde Park thinking that the party was already in place, regardless of their presence. This is analogous to how the self, although posited in a way that turns out to be fictional, might thereby be constructed in fact. So the positing of a self, that is in fact fictional, could end up in resulting in an actual self made up of unified psychological events[4].

If what the narrative is centred around isn't real, then we might also suspect that what it unifies cannot be real. However, the fact that Stephen Dedalus never

existed fails to imply that what is said to be Stephen Dedalus's history is equally non-existent, for the history in this case was Joyce's. The narrative reflects a human being's particular passage through time and space. What remains fictional is the status of a single simple mental self about which the entire narrative is true. Such a self is indeed a useful fiction. But couldn't these tales be the unified narrative of something non-fictional, such as a human being's conscious life for example? After all, the narrative reflects what a human was, is and in some senses will be. If the unified narrative or role is not itself fictional, but real, then this unified narrative or role, in conjunction with its physical realisers, could actually constitute a real self. For if "il n'y a pas de hors-texte" (Jacques Derrida in Dennett 1993: 411) is, as Dennett himself claims, close to the truth, then we should not ignore the unity and the boundaries of the text (the narrative, role or unified psychology) itself.

Just because what is posited as a self may well turn out to be a fiction does not imply that the self *is* a fiction. For what one takes to be a self needn't be what the self actually is. To conflate the conception and what it is the conception of is a mistake. The fact that our conception of the self may well be fictional does not mean that this fictional status will have to be transposed onto the self, should our self-conception be confirmed as fictional. There is surely a distinction between one's conception of something and that something itself – should it exist. Indeed, the so-called fictional status of the self may simply be indicative of the fact that one is misrepresenting something (or an aggregate of things) that might be thought of as a real self. (Although, in the case of selves, that misrepresentation might actually turn out to be a necessary ingredient in constructing the thing misrepresented.) For example, it might be common to conceive free choice as being somehow outside and impervious to any degree of environmental conditioning. However, if we were to discover beyond reasonable doubt, that acts of such accentuated randomness could never occur, we would not then be justified in claiming free-will to be non-existent. It seems that all we would be justified in claiming is that a *certain conception* of free-will is pure fiction.

It might be thought that it is the subject of the text (narrative or psychological properties) that we are inquiring after when inquiring about a self, and not the textual properties themselves, however unified they may be – not Hamlet the role but Hamlet the man. After all, aren't I the owner of record rather than that which is owned? But once we have accepted that such a self is most likely just a useful tool, and that nothing answers to it in the real world, maybe we should look elsewhere for the self. This conception of the self as a simple attractor of properties may well be useful, but useful for what? Creating a real self, perhaps. For the conception of the self as central owner is useful for establishing a unified psychology (or narrative) and this is a real effect of the fiction. But is it an effect that directly relates to selves? Dennett is himself confused, or at least unclear, about this issue. For instance, let us consider Dennett's closing comments on the possibility of immortality. He claims that if indeed we are centres of narrative gravity, then our life's continuing is reliant on the continuance of our particular narrative.

He claims that

> If what you are is that organisation of information that has structured your body's control system. [...] then you could in principle survive the death of your body as intact as a program can survive the destruction of the computer on which it was created and first run (Dennett 1993: 430)[5].

But this line of thought fails to distinguish the created self from the means by which it is created i.e., the posited self (think of Hamlet the non-existent man) from the structured unified narrative (think of Hamlet the role) that is thought to collect around it and be true of it. If what we are is, as Dennett puts it, "that organisation of information that has structured your body's control system," then we are not necessarily the principle of that organisation and likewise fictional. So although thinking of yourself as the centre of your narrative may well be a prerequisite for constructing a self, the fictional self at the centre of your narrative needn't be all that there is to being possessed of a self. For if I am the fictional posited self at the centre of my narrative, then I can't really survive anything. To survive something I would have to have existed in the first place. Dennett's claim that human beings construct selves in a manner analogous to beavers constructing dams and bower birds constructing bowers also seems to imply that the self is the unified narrative or role surrounding the fiction, as opposed to the purely fictional subject of that narrative.

One of the central problems of selfhood that my adaptation of Dennett's proposal might be thought to answer is that of continuity through time. There are no problems with how to account for a robust and enduring self, if indeed it is a fiction with such abilities written into its role. By this means the self can be believed to remain the same through any amount of seeming change, in much the same way as Peter Pan can be believed never to grow up. So we might reasonably claim that the fictional self can provide us with a conception of ourselves as continuous over time. It seems fair to say (in the absence of any other explanation) that only a fictional self can possibly provide the immutable and ever-present mental element to which all things in the narrative refer back, that which remains the same despite all the serpent-windings of the narrative. That is, unless we believe in Cartesian Egos, or the Bare Locus view[6], or indeed believe that there is a part of the brain that is the enduring self.

The narrative itself could be said to maintain continuity as a result of being collected around the same fiction. The representational posited self might enable the continuity and steady flow of the psychological narrative, in a similar way to how the dummy hare enables the form and directional momentum in a greyhound race, whereby collecting around a common fiction supplies a reality of common form to the elements that share in the same fiction. It is my conjecture that that common form, and the psychological components so formed (in conjunction with physical realisers e.g., a brain and a body) constitute a concrete self. It is my claim that the fictional self most likely provides the binding agent[7] to the beliefs, desires, reasons etc., which themselves constitute a real self as a placebo effect of the fictional self. For a real cure to be realised in cases of placebo effect the patient must have faith in the fact that they are already cured (or at least on the road to recovery). Similarly, for a real self to be constructed, we

must already believe that there is a self in place (i.e., this belief or conception being all there is to the fictional self). As we shall see in due course, this notion also applies and makes sense of abnormal cases such as multiple personality disorder (MPD).

3. Indeterminacy and Reality

I will now touch on the relation between determinacy and reality. I support the view that reality does not necessarily entail determinacy. This view is in direct opposition to Dennett's own; for Dennett believes that something's being indeterminate implies that it isn't real. Dennett allows of no vague or fuzzy existents; all that exists is, according to him, determinate[8].

George Graham (1999) claims that Dennett's whole rejection of the reality of the self takes the form of a sorites argument. Graham states Dennett's case as follows: (1) sufferers of MPD have more than one self. (2) It is impossible to decide upon just how many selves any one sufferer has. (3) There is no clear difference between how many selves a normal person has and how many a MPD sufferer may have (the distinction between normal cases and abnormal cases is fuzzy) (4) There is no clear fact of the matter how many selves a normal person has. (5) Therefore, selves are unreal in both normal and abnormal cases. I will briefly discuss Graham's response to this account of Dennett's argument, and by doing so offer an alternative to Graham's conclusion, while still supporting the realism that drives his answer.

Graham's initial response mirrors my own, for he too claims that a self's being indeterminate needn't deny it reality. But Graham ends up favouring an 'epistemic' response to this issue of first-person vagueness, as opposed to advocating an answer that talks in terms of degree. He claims that sufferers of MPD are selves, as are normal humans, we just don't know how many selves that might be. Prior to reaching this conclusion, Graham claims that one serious problem with the notion of fuzzy selves is that in introspective awareness I seem to be aware of "one definite self viz. me...there is no room for being more or less me" (Graham 1999: 62). However, I doubt whether self-consciousness does actually reveal such an unproblematic and definite oneness of self,[9] and even if it does on occasion seem that way to some I feel that my view offers a far more palatable explanation of why this might be the case. I suggest that feelings of concrete oneness, (such as Graham claims he has experienced) when engaged in introspection, could be said to be nothing more than the awareness of a fiction – the real self having no such pristine singularity. So the "gap" Graham points out, "between how I appear to myself (not fuzzy) and how I am supposed to be (fuzzy)" (Graham 1999: 62) could be filled or at least bridged, by accepting the role that the fictional self plays in the construction and maintenance of selfhood. Let us now look at Graham's case against the plausibility of fuzzy selves in greater detail.

Although Graham's 'strong ontic ignorance thesis' is perfectly able to accommodate the reality of indeterminate selves[10], he has two concerns about there ever being such selves. Firstly, Graham worries that accepting the self as indeterminate, "requires that there is no such thing as both seeming determinate and

being determinate" (Graham 1999: 62). In other words, because there is a normative distinction to be respected between a conception and what it is a conception of, then we have to allow not only for the possibility that the conception might not truly reflect its object, but also for the fact that it might. After all, "(p)erhaps I am as I appear" (Graham 1999: 62). But nobody is denying that this is a possibility; it's just that given what we know about the brain and how it works, it is not the most likely of the possibilities.

Graham's second concern is that the reality of fuzzy selves would result in a gulf between our phenomenology of self and how selves really are. That if I am as I appear to be in introspective awareness, then selves are not indeterminate. He writes:

> If I am fuzzy or indeterminate, then it makes sense to say that I can become more and more self. However I don't seem like I can become more and more myself. I have only to reflect just a little to see that if I'm a real self and not just a fictional entity, I'm something definite already. That something is me. I do not seem to be more or less me (Graham 1999: 62).

If I am right and the self does have a fictional element, then the sense of oneself as simple and definite could be explained by this element. But this aside, the passage quoted above makes some unwarranted assumptions about the fuzzy conception of selves. The fact that I may be fuzzy or indeterminate does not mean that I can become more and more myself. For this would seem to suggest that there was something over and above my body, my brain, and certain mental and physical events, that contained my essence and to which I could progressively assimilate. If I could become more and more myself then it is likely to follow from this that I could become less and less myself. If this could happen, then I could get to a stage where I am nothing like me but nevertheless still remain me. This again presupposes some determinate archetype of me that I could both fall short of and come close to being like. But there is no underlying 'I' to become more or less itself. There is only a body, a brain and a more or less connected and continuous set of psychological events. If I lose a great deal of the connectedness and continuity that exist between my psychological states I do not become less and less me, just less and less connected and continuous with my former self or selves.

Graham's concern, with what he calls "New Fictionalism" about the self, is that such a position vastly overestimates what we actually know about the self. In other words, we cannot confidently claim the self as a fiction when we don't really have a tight grasp on what the self is. I think we do in fact have a relatively comprehensive conception of what selves are (or at least what criteria they satisfy) and it just so happens that certain aspects of that conception are best explained by incorporating an element of fiction into what is primarily real and encounterable. We should stick to what we know about ourselves and define selves in terms of that knowledge, if indeed there is room for them. What we should not do, is follow Dennett and conclude that there is no self, just because a particular conception of the self has no objective status. Let us now consider exactly what it is that is fictional about the self and the context in which Dennett hopes to make his point.

The prime target for Dennett's notion of the self is what might be called a Cartesian one, but more specifically it is anyone who thinks of the self as a "mind-pearl", a "brain-pearl", or a simple and constant central controller[11]. (Modern proponents of such a view would be Roderick. M. Chisholm and Richard Swinburne, both of whom follow in the footsteps of Butler and Reid). However, the denial of such a position is hardly controversial, in fact in philosophical circles it might even be regarded as commonplace. But Dennett thinks that refuting the mind-pearl conception of the self (or rather explaining it as fiction) is sufficient to leave his own position as the most likely (or only) alternative when, truth is, he faces a more prevalent and indeed stronger opponent in the Psychological Reductionist[12]. Such a theorist agrees that the self is not simple, but in claiming selves to be composite and indeterminate does not thereby deny the reality of selves. Such a view of the self is definitely not encompassed by the term 'mind-pearl'. With regard to the theorising itself, Dennett would (and in fact does) agree with a lot that the Psychological Reductionist has to say about the self. But Dennett would say the Psychological Reductionist has made a mistake in failing to realise that such observations of indeterminacy either imply that the self has fictional status, or that a unified psychology doesn't qualify as a self.

Again we see that Dennett's criterion of reality, along with his insistence on eliminativism being the right approach toward the self as opposed to reductionism, result in a rather extreme and narrow account. Dennett maintains that only determinate things exist; so by then proving the self to be indeterminate, he thereby supposes that he has also proved the self to be unreal. Any reality the constituents of the narrative might have, they do not, for Dennett, have them as a self. Does this mean that Dennett is forced to conclude that things such as clouds and heaps don't exist either? Not necessarily, because we must take note of what he says in Appendix A.

> The question of whether to treat part of the heterophenomenological world of a subject as a useful fiction rather than a somewhat strained truth is not always a question that deserves much attention. Are mental images real? There are real structures in people's brains that are rather like images – are they the mental images you're asking about? If so, then yes; if not, then no (Dennett 1993: 459).

So in the cases of 'heat = molecular motion' or 'clouds = masses of condensed water vapour' we do not have to claim that heat and clouds are fictional, and this is due to the fact that they share reference with actual physical goings-on. They are non-illusive re-descriptions. The *self*, however, does not fall into this category, according to Dennett, because there is nothing sufficiently similar in us that our notion of the self could be thought to represent. There is no physical thing in the brain to which the term 'self' could legitimately be said to refer, and no simple and impredicable attractor of properties. Therefore the self is a fiction as opposed to a "strained truth".

No one painting or artist constituted the Impressionist movement. We cannot be absolutely sure where or when it started[13], although we can pick out acknowledged masters and a rough period of seventy years as its lifespan. We can delve into the history of the movement; look at inspirations, attitudes, progressions and

changes. But there is no doubt that there was an art movement that we now call 'Impressionism'. We can interpret it in many different ways, include other possible contributors etc., but all such reanalysis will have to remain faithful to the stylistic facts of the movement or else risk inaccuracy. Similarly, somebody's life may be reinterpreted by themselves or others, other motives included etc., but an accurate analysis will have to respect the facts of the life in question. Also, artists didn't first posit something called 'Impressionism' and then go about working on it, rather it materialised as a result of a collective artistic endeavour. Similarly, we do not initially go about with a mind to constructing a self; it comes about as a result of a human-being's biology and its worldly interactions. The positing of a fictional self may be an invaluable process in the construction of a real self, likewise an overriding concern for the primacy of the act of beholding and the importance of colour are seen as important considerations for any Impressionist artist. And just as Impressionism is made up of real things e.g., paintings and artists, so too is the self e.g., brains (or functional equivalents), bodies, and psychological and physical events. Hopefully, this analogy serves as an example of how the construction of something real can be initially blind and that real things can still be somewhat indeterminate in duration and location.

But we are still faced with the question whether our fictional selves are represented as determinate or indeterminate. I mention this because Dennett starts by suggesting that our self-conception is of a unified and robust agent, but then praises Parfit for correctly stating the nature of the self as analogous to a club and therefore indeterminate. So if the self is just the centre of narrative gravity, then what is it that Parfit is right about? After all, if there is nothing to the self but the conception of the self and that conception doesn't portray the self as indeterminate, then what else is there for Parfit to be right about? Let me lay this point out more clearly so as to get to grips with what Dennett means by insinuating that the self is both conceived as determinate and really indeterminate, despite not really being anything. Consider these claims made by Dennett:

(1) "(T)he self is 'just' the Center of Narrative Gravity".
(2) The centre of narrative gravity is a human being's "self representation".
(3) There is nothing to the self but the way that the brain represents it.
(4) Having a self involves believing a fictional device to be truly representa tional.
(5) The centre of narrative gravity is fictional; there is only the conception of it – its being posited.
(6) As a fiction (an abstractum) the self is useful – it is "a unified agent" "the owner of record", "a remarkably robust and almost tangible attractor of properties".
(7) Parfit is right, the self is really gappy and indeterminate.

It appears that (1) through to (6) deny the possibility of (7) being the case. For if there is nothing to the self but one's conception of it, and that conception is of a certain sort *x*, then how can there be another sort *y* (completely unlike *x*) that it is *really* like? The answer lies in the fact that Dennett is talking about different things without making it clear that he is doing so. It seems that what Den-

nett wants to say is that the simple posited self is fictional and so not real, and that the constructed self (the unified narrative or psychology) is indeterminate, and so also not real, a unified collection of psychological states being too different from that commonly represented or thought of as a self to legitimately qualify as a self. The term 'self', he would argue, is not just a mere re-description of such a collection.

Dennett claims that outside the confines of a fictional text "(t)here is simply no fact of the matter" (Dennett 1993: 278). This implies that there is some fact of the matter within the confines of the text; there is certainly some fact of the matter within one's own narrative. Outside one's own history, as outside a fictional text, there are indeed no facts of the matter. Just as there is no fact of the matter whether, had I lived in 14th Century Italy I would have fallen in love with Dante's Beatrice, or had I been sired by Minnesota Fats I would have been a famous pool player, there is no fact of the matter about whether Dimitri, Alyosha and Alexi all had moles on their left shoulder blades, or whether Alice played hopscotch in the period between visiting Wonderland and venturing through the looking glass. However, inside one's narrative history the facts have a different status to those concerning some fictional text. The following statements are true, in that they reflect events that occurred in the real world: "I caught the bus yesterday", "I recalled the rooftop wonders of Casa Batlló yesterday". But these facts do not seem to be directly comparable to facts about a fictional text, such as: "In the Sherlock Holmes stories Holmes often patronises Watson", or "In *The Speckled Band* an Indian snake causes the death of Helen's sister". In the latter, we require the relevant prefixes ("In the Sherlock Holmes stories..." "In *The Speckled Band*...") in order to qualify the facts as relevant only to some particular fictional text or other.

While a self to which all the events in one's life are attributed may be as fictional as Sherlock Holmes, the events attributed to a self by a brain are not as fictional as the events attributed to Sherlock Holmes. The events that are attributed to me are events that took place in the spatio-temporal framework of the real world: they were events in the mental and physical history of some part of reality i.e., a particular brain and a particular body, traversing a particular spatio-temporal path. Whereas the events that are attributed to Sherlock Holmes in Conan Doyle's stories simply never occurred.

Dennett's response to the above might be something like the following: "A body is not a self, a brain is not a self, and a collection of physical and mental events is not a self. It is a self that I claim is a fiction and these things are not selves, nor do they jointly compose a self. It is what supposedly has all these things that I claim to be a fiction and what a self is thought to be". Given that these elements do in fact fulfil the Psychological Reductionist's criterion for what a self is thought to be, a criterion with which Dennett claims he is sympathetic, I believe Dennett is wrong. Dennett would, no doubt, appeal to selves being both simple and somehow impredicable in the sense favoured by the Bare Locus view, or occupying a specific part of the brain. In this sense, of course, a brain, a body and a series of physical and mental events would not make up a self, and such simple selves would indeed, it seems, be pure (if not evolutionarily beneficial) fictions. But is it necessary that something be simple and impredica-

ble, or a specific part of the brain, if it is to qualify as a self? It seems that these are just ways of characterising selves, but not the only ways. And doesn't the fact that there is no part of reality to answer to either of these descriptions not only suggest that their subjects may be fictional, but also that the descriptions themselves might be wrong?

4. Conclusion: Fledgling Selves and the Wandering Boundaries of Fictionalism

I now wish to discuss Dennett's comments on split-brain patients. In this discussion I will not only reinforce the point that Dennett doesn't consider the self as merely a *centre* of narrative gravity (analogous to the man Hamlet) but will also argue that his implicit broadening of the notion of a fictional self means that his notion of the self also constitutes a somewhat surreptitious attack on the Psychological Reductionist theory of the self.

Dennett's comments on split-brain patients imply that there are gradations of selfhood, dependent on the richness of the narrative. This is a further indication that he includes the narrative or psychological states that surround the centre of narrative gravity within his concept of a self. For if Dennett really considered the self to be a *centre* of narrative gravity, then he couldn't talk of gradations of selfhood because such a self is simple and so doesn't allow of gradation.

According to Dennett, a self must last a certain length of time and accrue a certain amount of autobiography if it is to be considered "fully fledged". The sort of self that fails to satisfy these criteria might reasonably be termed a 'fledgling self'. Dennett suggests that such a fledgling self is temporarily present in the right-hemisphere of a split-brain patient. He claims that

> For brief periods during carefully devised experimental procedures, a few of these patients bifurcate in their response to a predicament, temporarily creating a second centre of narrative gravity. A few effects of the bifurcation may linger on indefinitely in mutually inaccessible memory traces, but aside from these actually quite primitive traces of the bifurcation, the life of a second rudimentary self lasts a few minutes at most, not much time to accrue the sort of autobiography of which fully fledged selves are made (Dennett 1993: 425).

So a fledgling self such as this fails to be a fully-fledged self, not because there is no centre of narrative gravity, but because there is not sufficient time for this centre to accumulate a sufficient narrative. This actually implies that the centre of narrative gravity is nothing but a fledgling self, and so merely a prerequisite for the formation of a fully-fledged self, and so not all there is to being a self[14]. Again we see the regularity with which Dennett moves away from his position that the self is an abstract subject, towards the idea that a self is constructed from the narrative; as it is "narrative richness... that constitutes a "fully fledged" self" (Dennett 1993: 426) and not the centre of that narrative.

But given this, Dennett is in a far trickier and far less intuitively appealing position. For once we consider what it is that constitutes this fully-fledged self, we find that its status as pure (non-existent) abstractum is harder to accept, at

least without a convincing argument, which Dennett neglects to supply. So despite Dennett claiming the mind-pearl conception of the self as his target, he also wages an attack on the Psychological Reductionist view of the self by furtively (and perhaps even unwittingly) including that conception of the self (the narrative itself) within the scope of his attack. He continues to switch from the claim that the self is a centre of narrative gravity, to the claim that the narrative is the self (or at least an integral part of it) and then just assumes that the narrative itself is also pure fiction. What Dennett calls "narrative" is what the Psychological Reductionist believes to be the most important element of the self – the element of the self that matters – i.e., the connectedness and continuity of psychological events, that which is predicated of the subject. All the fluctuations and continuities of a human being's history are claimed by some to have a unified agent that resides behind that human being's deeds and manifestations. Dennett is right in claiming that such an agent is fictional, but he is not justified in transferring that conclusion to what that agent was supposedly a centre of i.e., the fluctuations and continuities of a human being's psychological history. Similarly, we can explain the swaying and toppling of a chair with an abstraction (the chair's centre of gravity) but do not end up transposing that same ontological status to the chair itself; there is no such man as Hamlet, but that is not to say that the role Hamlet is equally non-existent.

The view of the self that Dennett puts forward turns out to be rather counter-intuitive and somewhat confused. I believe my account of a real, indeterminate and functional self that results from a fictional conception of a self retains the prima facie appeal of Dennett's position, while avoiding the troublesome implications and confusions of that position.

Endnotes

[1] It has been supported at one time or another by the likes of Hume (who merely claimed its *simple status* as fictional), Nietzsche, Russell, and Wittgenstein.

[2] For despite commenting on the autobiographical nature of the fiction, Dennett fails to realise the full implications of a fiction with autobiographical status.

[3] The question as to whether selves could exist as abstract entities shall not be dealt with here. I shall limit myself to talking of real (or concrete) selves, while leaving open the possibility that selves might exist in some abstract form.

[4] Dennett writes: "(aren't all fictional selves dependent for their very creation on the existence of real selves? It may seem so, but I will argue that this is an illusion." (Dennett, 1992) I suggest that maybe it's the other way around i.e., maybe the formation of real selves is dependent on the existence of fictional selves.

[5] Although I do not have the room to go into this here, this is something I believe to be possible.

[6] The Bare Locus View is what Derek Parfit has called the 'Further Fact View', according to which facts about identity are facts over and above any facts that are about the constitutive properties of the objects in question. But facts about persons should surely be facts about things and events in the world and somehow constitute the way things are, or the way they could be. What then is the Bare Locus View supposedly true of? The simple and most obvious answer is that it is

true of persons, rather than being true of the properties of persons i.e., it is true of that which has personal predicates accurately ascribed to it (a bare locus or subject) as opposed to being true of the conglomerate of properties which those predicates pick out. But whatever logical subjects of personal predicates are, you and I are not such things. As Kathleen Wilkes puts it: «The moral seems to be that if one has faith in what I have called a 'self-spark', then one will be....open to the challenge to say something about this spark. And since, I claim, the spark is a matter of faith rather than of reason, then it is a claim which cannot be rationally defended: a claim with which philosophy can do nothing» (Wilkes 1991: 243).

[7] This 'binding agent' is real in the sense that it achieves the purpose of cohesion, fictional in the sense that there is nothing that is referent to one's conception of it.

[8] Unfortunately I do not have space to discuss this matter in anything like sufficient detail. But hopefully I can succeed in weakening some of the arguments of those that question the plausibility of vague selves being real.

[9] Consider Hume's own response to talk of such an awareness "If anyone, upon serious and unprejudiced reflection, thinks he has a different notion of himself, I must confess I can reason no longer with him". However, my doubts do not, I feel, have such a dramatic result, for my notion of the self can account for that "different notion" without having to buy into the reality of a simple self.

[10] It is able to do so because the limited nature of our knowledge of selves does not provide us with a concrete answer as to whether we are indeterminate or determinate entities.

[11] In the way in which our identity is claimed to be simple and beyond empirical evidence, the Bare Locus view can also be regarded as a target of Dennett's theory.

[12] Psychological Reductionism is a broadly Parfitian approach to persons and their persistence conditions. Psychological Reductionism maintains that the existence of a person consists in the existence of a brain, a body, *or functional equivalents* and a series of interrelated mental and physical events (see Derek Parfit: 1984).

[13] Although we do know that the term 'impressionistic' was coined by a critic, who meant it as a criticism of a particular way of painting – a way of painting that we associate with Impressionism.

[14] This is in fact very similar to the line I am taking toward the self apart from the fact that Dennett says that even a "fully fledged" self is fictional.

References

Brody, B.A. (1980). *Identity and Essence*. Princeton University Press.

Dennett, D. (1978). "Where Am I". In *Brainstorms: Philosophical Essays on Mind and Psychology*. Bradford Books.

Dennett, D. (1992). "The Self as a Center of Narrative Gravità". In F.Kessel, P.Cole, D. Johnson (Eds.) *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ, Erlbaum.

Dennett, D. (1993). *Consciousness Explained*. Penguin Books.

Graham, G. (1999). Self-Consciousness, Psychopathology and Realism About the Self. *Anthropology and Philosophy*, 3, 57-66.

Graham, G. (1999). Fuzzy Fault Lines: Selves in Multiple Personality Disorder. *Philosophical Explorations*, 3, 159-174.

Parfit, D. (1971). The Importance of Self-Identity. *The Journal of Philosophy*, 68, 667-690.

Parfit, D. (1971). "Personal Identity". In Glover (1976).

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Wilkes, K.V. (1988). *Real People*. Oxford University Press.

Wilkes, K.V. (1991). "How Many Selves Make Me". In Cockburn (1991).

Finito di stampare nel mese di giugno 2008
presso la Stampa Editoriale srl
STRADA STATALE 7/BIS - ZONA INDUSTRIALE DI AVELLINO
83030 – VANOCALZATI (AV)
TEL. 0825/62.69.66 - LAX 0825/61.08.88