UNIVERSITÀ DI PISA

DIPARTIMENTO DI FILOLOGIA, LETTERATURA E LINGUISTICA


SCUOLA DI DOTTORATO IN DISCIPLINE UMANISTICHE

PROGRAMMA DI RICERCA IN LINGUISTICA GENERALE, STORICA, APPLICATA, COMPUTAZIONALE E DELLE LINGUE MODERNE

L-LIN/01

*Tesi di Dottorato*

# LANGUAGE AND AFFORDANCES

*Candidata*

IRENE DE FELICE

*Tutori*

PROF.SSA GIOVANNA MAROTTA

DOTT.SSA NICOLETTA CALZOLARI

DOTT.SSA MONICA MONACHINI

*Presidente*

PROF. ALESSANDRO LENCI

TRIENNIO 2012-2014

# Index

# Introduction

> An affordance is neither an objective property nor a subjective property; or it is both if you like. An affordance cuts across the dichotomy of subjective-objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer (Gibson 1979: 129).

The concept of affordance was first introduced by James Gibson in the late 1970's within the field of the ecological psychology. Affordances are defined as possibilities for action that the environment offers to living beings. The basic idea of this theory is extraordinarily intuitive: when we perceive the world surrounding us and the objects in it, we actually perceive potential actions. For instance, when we look at a handled cup full of tea, we perceive the possibilities to grasp the object by its handle and to drink the tea.

With the theory of affordances, Gibson highlights the close connection existing between perception and action. More recently, a growing body of literature in the cognitive neurosciences has provided strong empirical evidence that the sensory and the motor systems are strictly interconnected and interfere with one another in a complex way. In

particular, a fundamental and fascinating contribution to the theory of affordances comes from the behavioural and neurophysiological research conducted on objects' manipulation and grasping, which demonstrates that the same cortical areas are activated both when an agent grasps and manipulates an object, and when the agent simply observes the object, even without any intention to act (Chao and Martin 2000; Grèzes and Decety 2002; Grèzes, Armony et al. 2003). Such findings point to the concrete existence, at least in the specific field of grasping, of affordances, intended as the simulations of possible actions automatically triggered by the perception of visually presented objects (Tucker and Ellis 1998). For instance, when we observe a handled cup in certain conditions, our brain automatically constructs a mental simulation of the actions necessary to grasp it, and recruits the same neurons that would be active during a real grasp of the cup. Experiments also show that sensory-motor responses to visual stimuli presentation are strictly related to, and modulated by, the properties of the perceived objects (e.g. their location, orientation, semantic type, constituency, shape, and dimension). For example, if the handle of the cup is broken, the activation of the motor system after object perception is much weaker (Buccino et al. 2009).

In light of this evidence, the main research question that guides this work is the following: does language reflect affordances? In other words, is language sensitive to the same variables that modulate sensory-motor responses to visually presented objects? Such a question leads to a hitherto unexplored area of research, to which this thesis intends to contribute.

This very broad and intriguing issue will be tackled within the specific domain of grasping, in order to allow a more easy comparison with the existing works. In order to explore the relation between language and affordances, a speech corpus of grasp descriptions has been collected and analysed at different levels. The primary purpose of this study is to investigate whether the way in which people linguistically describe the grasping of a visually presented object is modulated by the same factors that are responsible for modulating brain activity within the domain of grasping (i.e. with relation to the affordance of "graspability").

The work is structured as follows. In Chapter I, we will outline a brief history of the concept of affordance, since its first emergence within the field of the ecological psychology to its later development in other research domains. In particular, a huge number of behavioural and neurophysiological studies conducted on grasping and manipulation (mostly based on visual stimuli presentation) have a major impact on the theory of affordances and constitute an important part of the background of this study. Therefore, they are well worth

presenting and discussing, although they are not strictly related to linguistics. Nevertheless, the last part of the chapter will be devoted to some recent findings from neurolinguistic and psycholinguistic research, which offers insight into the complex relation between perception, action, cognition and language.

Chapter II illustrates the action description task explicitly designed to investigate linguistic reflexes of the affordance of graspability. In this experiment, subjects are visually presented with 42 pictures representing graspable and/or manipulable entities and are asked to describe how they would grasp such entities. The system and the specifications adopted to transcribe the interviews are also explained.

The two core chapters contain the analysis of the linguistic descriptions of grasps. In particular, Chapter III considers the distribution of explicit references to the effector of the grasp (such as the mention of the hand, which reflects a focus of attention on the agent involved in the action) and the target of the grasp (such as an object's handle, which instead reflects a focus on the perceived object). A more in-depth analysis of the linguistic data is presented in Chapter IV, in which the lexical words used by informants to refer to the effector or the target of the grasp are extracted from transcripts and classified according to a set of semantic classes. The purpose of these two studies is to highlight whether the difference (in terms of constituency, shape, orientation) between the objects-stimuli adopted in the experiment corresponds to a difference in the linguistic production of informants, either in terms of references to the effector or the target of the grasp, or in terms of lexical choices used to describe the action. The results of these two complementary analyses are discussed in light of the behavioural and neurophysiological findings presented in Chapter I.

Chapter V describes the methodology and the results of a property generation (or feature-listing) task, conducted on a part of the stimuli adopted in the previous experiment. Whereas in the action description task stimuli consist of images of objects visually presented to informants (e.g. the picture of a jug), in this property generation task they consist of the written form of the words denoting the same objects (e.g. the word *brocca*, "jug"). Informants are asked to list a series of features that they consider relevant in order to describe the meaning of these linguistic items. The purpose of this second experiment is to establish a comparison between the explicit mentions of objects' parts (meronyms) produced in the action description task to indicate the target of the grasp (such as *manico*, "handle") and those produced in the feature-listing task.

Finally, Chapter VI presents some possible applications and further developments of the research conducted. In particular, two case studies are proposed. The first one, more relevant to computational linguistics, explores the possibility to annotate information related to affordances in a corpus. The second one, more relevant to cognitive sciences, integrates gesture annotation in a small part of the transcripts derived from grasp descriptions and offers an interesting view on the effects of objects' orientation and informants' hand dominance on the kind of grasp described. Other possible applications of the results obtained are discussed in light of the general purpose of ModelAct, the project within which the research illustrated in this thesis has been carried out. The main objective of the ModelAct project is to propose a model of the human categorization of action, in terms of both linguistic and cognitive encoding. To this aim, the project exploits the ImagAct ontology (that is presented in Chapter II), in which language-independent action categories are identified and represented as prototypical scenes. The primary goal of ModelAct is to go beyond the identification of action concepts and provide a more formal definition (i.e. a model) for them, which could prove useful in natural language processing and human machine interaction.

The last chapter summarises the main results of the entire work and adds some further general remarks on the relevance of the theory of affordances for modern disciplines.

# Chapter I

# Affordances

In this chapter, a brief history of the concept of affordances will be first provided, since its first emergence within the ecological psychology, to its later development in other research fields (§1.1, §1.2). The second part of the chapter will focus on affordances relative to objects' manipulation and grasping: this will serve as a basis for the following chapters of the thesis, centred on these kinds of events. We will review recent contributions from neurophysiological, neuropsychological and behavioural research that provide convincing arguments for defining these specific affordances as motor representation elicited by visually presented objects (§1.3) or words (§1.4). Finally, some conclusive remarks will be drawn (§1.5).

## 1.1 Gibson and the ecological psychology

In the late seventies, James Gibson introduced, in the field of ecological psychology, the concept of affordances (Gibson 1977; 1979). The author, in a well known passage, writes:

> The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, the noun

> affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment (Gibson 1979: 127).

As Gibson claims, he was the first to introduce the term *affordance*, referring to the possibilities for action that an environment offers living beings. For instance, if an agent is faced with a flat surface that comes up to about his knees, he is offered the chance to sit down: but knee height of a child is not the same as knee height of an adult.[1] It is evident, then, that the possibilities for action afforded by objects also depend on the agent who perceives them: «affordances are properties of things taken with reference to the observer» (*ibid.*: 143).

Gibson introduced the concept of affordances within the framework of the ecological psychology, whose main assumption is that the behaviour of living beings is anchored to the environment in which they are set and it is not possible to leave apart this "external" information, the setting in which every event takes place, from the study of behaviours. Perception obviously plays a central role in this theory: it is through the perceptual abilities of animals, as well as through the existence of perceptible features in environment, that animals may establish a relation with the environment and may detect affordances. This animal-environment system may be defined as a niche: «in ecology, a niche is a setting of environmental features that are suitable for an animal, into which it fits metaphorically» (*ibid.*: 129).

An ecological theory of perception, thus, assumes that perception is directly grounded in ambient and directed towards every kind of available information in there. In this sense, external information provided by the environment, and so direct perception, is more relevant than any other kind of indirect perception or internal sensation (as stated by Mace 1977, «ask not what's inside your head, but what your head's inside of»). From Gibson's point of view, direct perception is not a passive process of visual perception of objects *per se*, nor does it require high-level processes such as reasoning about object properties:

[1] «If a surface of support [...] is also knee-high above the ground, it affords sitting on […] Knee-high for a child is not the same as knee-high for an adult, so the affordance is relative to the size of an individual» (Gibson 1979: 128).

> In the realm of manipulation, for example, a person seeing an object would not necessarily only perceive colours, shapes and so on, but first and foremost also directly perceive the object's "graspability", "liftability" and so on (Thill et al. 2013: 492).

It is now possible to understand why Gibson uses the term affordances in a very broad sense, referring to almost every kind of possibility for action agents may find in the world (others will redefine later the limits of the concept): affordances comprise *every possibility for action* that living beings are able to seize in the environment, detecting any kind of directly perceivable external information. For example, humans may identify affordances also in other humans:

> What other persons afford, comprises the whole realm of social significance for human beings. We pay the closest attention to the optical and acoustic information that specifies what the other person is, invites, threatens, and does (Gibson 1979: 128).

In this sense, Gibson can adopt the term affordance with regards to human-human interaction because our behaviour is always deeply influenced by the others' behaviour and a sort of reaction to it.

### 1.1.1 An inherently relational concept

Since its first definition, the concept of affordance is presented as an inherently relational one. It is clear that, in Gibson's theory, affordances are different and unique for each agent, since they are not simply related only to visually perceivable properties of objects («an affordance cannot be measured as we measure in physics», *ibid.*: 128); rather, they reside in the possible ways in which living beings can interact with objects.

A number of psychological studies explored this issue and proved that the judgement about the ability to perform particular actions depends on both the physical characteristics of agents and the perceivable characteristics of objects. For instance, Warren (1984) assumes that in order to act agents must be capable of perceiving the relationship between the environmental properties and the properties of their own action system (note that the properties he refers to may rely on geometric variables, e.g. size, dimension, as well as on kinetic variables, e.g. mass, force, friction, elasticity, work, relevant to metabolic energy

exchanges). He defines a specific set of values of the animal and environmental properties relevant to a given activity as a dynamic animal-environment fit. Focussing on the action of stair climbing, the metabolic efficiency of this action would thus be determined by the fit between the properties of a given stairway and a given climber. What is interesting is that in his study he demonstrates that the "best fit" is always reached at a constant point, according to a body-scaled metric. In his experiments, he asked two groups of people (one composed of tall and the other of short people) to judge if the stairs they were seeing in a photograph were climbable, and also to express a "confiance" judgement. Answers show that as properties vary (namely the dimensions of the riser, the leg length), judgements change as well. In particular, the author individuates critical points, after which people started to deem the stair unclimbable (e.g. when riser height increases and reaches a height that cannot afford bipedal climbing), as well as optimal points (that best fit affordances).

To clarify this idea, it will be useful to quote a passage from one of Warren's later works:

> As the fit is varied, optimal points in the ecosystem may emerge for preferred states at which a given action is most comfortable or efficient, and critical points will emerge at which the limits on an action are reached and a phase transition to a qualitatively different action occurs (Warren and Whang 1987: 371).

The analysis of data from the two groups, as was expected, gave different results about the exact determination of critical and optimal points: tall people considered as climbable stairs that for short people were unclimbable. Nevertheless, optimal/critical points turned out to be invariant proportions of agents' leg length and riser height. The category boundaries emergent from this study are constant over changes in scale (short/tall agents, different measures of the riser), thus demonstrating that perception is anchored in the biomechanics of activity. Furthermore, perceptual preferences matched with optimal preferences: looking at stairways, both short and tall people preferred a riser height that coincided with the optimal size.

The methodology adopted in Warren's experiments and the results obtained, in particular the principle of intrinsic measurement and the dynamics of critical/optimal points, lead to quite general considerations that impact on the theory proposed by Gibson and confirm many of his intuitions.

In addition to this study, many other works were conducted exploring other affordances, such as that of passing through an opening (Warren and Whang 1987) or crossing gaps (Mark et al. 1999; see also Chemero et al. 2003), that highlight the strong relation existing between, on the on hand, the possibility for a given agent to perform a given action, and on the other, the physical characteristics of the agent, together with the characteristics of the environment in which the event takes place (see also Mark 1987; Mark et al. 1990).

### 1.1.2 *Where* are affordances?

Strictly linked to the relativity of the concept is the question about "where" affordances are. According to Gibson, as already stated, affordances may be conceived as «properties of things *taken with reference to an observer* but not properties of the *experiences of the observer*» (Gibson 1979: 137, original emphasis). This means that affordances do exist independently from the existence of a perceiver and are true properties of entities: they «are not created in the act of perception» (Michaels 2003: 136). However, to be called affordances, objectual properties have to interact with the properties of a specific perceiver, so that an activity can be supported:

> An affordance, as I said, points two ways, to the environment and to the observer. So does the information to specify an affordance. But this does not in the least imply separate realms of consciousness and matter, a psychophysical dualism. It says only that the information to specify the utilities of the environment is accompanied by information to specify the observer himself, his body, legs, hands, and mouth. This is only to reemphasize that exteroception is accompanied by proprioception - that to perceive the world is to coperceive oneself (Gibson 1979: 141).

This point turns out to be a bit confusing and, for a reader, affordances may look like «impossible, ghostly entities» (Chemero 2009: 136). Thus, in many post-Gibsonian works, affordances are fundamentally seen as environmental properties; but the central idea of environment-agent mutuality is not lost, because these properties, to be true affordances, must be complemented by agents' *effectivities* (Turvey et al. 1981; Turvey 1992; Shaw et al. 1982), *abilities* (Greeno 1994) or *aptitudes* (Snow 1992). Even when different labels are used

to distinguish the two poles of the same relation, the relationality of the concept is still a strong point of these theories, as it appears clearly from the following quotation:

> Affordances and effectivities […] are complementary properties of animal and environment. In this view, affordances are properties of the environment that permit an animal to execute certain action in, whereas effectivities are the properties of the animal that allow that action to take place in the environment (Michaels 2003: 139).

In more recent years, Chemero (2001; 2003; 2009) tries to solve this difficulty intending an affordance not as the relationship between two distinct entities, i.e. the properties of the environment, on the one side, and the properties of the agent, on the other, but as «features of whole situations», provided that «animals are, of course, usually crucial parts of these whole situations, so perceiving something about the whole situation cannot always be just perceiving something about the environment, divorced from the animal» (Chemero 2003: 185).

### 1.1.3 Affordances as distinctive objectual features

Another important issue related to this debate is that affordances were also often considered as invariant properties of objects that are able to distinguish and characterise them. For this reason, they somehow resemble other objects' properties, such as colour, shape, dimension etc.; but according to Gibson, affordance-related properties are for humans even more salient than other types of perceivable features:

> Orthodox psychology asserts that we perceive these objects insofar as we discriminate their properties or qualities […] I now suggest that what we perceive when we look at objects are their affordances, not their qualities. […] The affordance of an object is what the infant begins by noticing. The meaning is observed before the substance and surface, the colour and form, are seen as such. An affordance is an invariant combination of variables (Gibson 1979: 134).

This is because the act of perception, according to the author (*ibid.*: 135), is governed by the principle of economy: agents need not perceive all the properties of an object to recognise it and distinguish it from other things (perhaps it would be impossible to do so);

agents only notice the minimum of distinctive features of objects. This topic was further investigated by Gibson's wife, Eleanor Gibson, that in following years explored the concept of affordances within the field of developmental psychology (she devoted a book to this argument, written with A. Pick in 2000: *An ecological approach to perceptual learning and development*). In Gibson (2000b), she studied the exploratory activity of children and suggested that affordances play a crucial role in their development: perceptual learning passes through the discovery of distinctive features and invariant properties of objects.

### 1.1.4 The role of visual perception

In the passage quoted from James Gibson in the last paragraph, there is another interesting point (emphasis added): «what we perceive *when we look* at objects are their affordances, not their qualities». This mention allows us to introduce one central point of the whole theory of affordances, not only as it appears in Gibson's works but also in much of its later development, i.e. the fact that perception is intended almost always with reference to *visual* perception (as highlighted by Marotta 2013: 14, this is one of the weakest point of Gibson's theory of perception). As a proof of this, for instance, Gibson considers necessary, in order to perceive affordances, sufficiently good light conditions: «the central question for the theory of affordances is not whether they exist and are real but whether information is available in ambient light for perceiving them» (Gibson 1979: 140). However, the same author refers to affordances also as *every* possibility for action offered by the environment to living beings; as already noticed, from his point of view also humans afford behaviour to other humans. But Gibson did not examine in depth these aspects, so that the concept of affordance itself seems entirely developed within the framework of a new theory for perception, that since the title of the famous 1979 book is restricted to the domain of sight (*The ecological approach to visual perception*). It is indubitable that visual perception is probably the first, most immediate and most informative way in which living beings discover the world and detect the possibilities for action it offers them, but there is also something else:

> When we grab a banana, our hands experience the texture of the banana peel, the ridges along the peel, the smooth extensions between the ridges, and the rougher edges where the banana connects with other bananas into a bunch. These haptic affordances are coordinated with visual affordances such as a perception of the yellow and brown colors

> of the banana and its curving shape. When we hold or throw a banana, we appreciate its weight and balance. An overripe banana can assault us with its pungent smell. When we peel a banana, we encounter still further affordances involving the action of peeling, as well as the peel itself. With the peel removed, we can access new affordances from the meat of the banana. When we eat a banana, our whole body becomes involved in chewing, swallowing, and digestion. All of these affordances in vision, smell, taste, touch, skeletal postures, haptic actions, and even locomotion are provided by a single object that we categorize as a "banana" (MacWhinney 1999: 218).

This passage well describes how many different possibilities for action a single object may afford, and not all of them are stimulated by visual perception. But here, MacWhinney conceives affordances in a very broad sense, as «sensations that we experience when we interact with individual objects» (*ibid.*), whereas in most theories, as in our opinion, affordances are better considered as «preconditions for action» (Greeno 1994: 340) and must be strictly related to action (not only to sensations). We would not say that, in normal conditions, the yellow colour of a banana is an affordance (which is the action afforded by the yellow colour of a fruit?).

## 1.2 Other perspectives on affordances

### 1.2.1 Design and Technology

After the first studies in the field of psychology, the original meaning of affordances, as outlined by Gibson, was partly changed and adapted to other fields. In particular, the idea that objects provide direct information on how they are supposed to be used can help planning and designing easily usable artefacts that suggest relevant actions in an immediate way.

For the design, the most notable contribution to the theory of affordances probably comes from Donald Norman, who is convinced that (1988: 123) «affordances are of little use if they are not visible to the users. Hence, the art of the designer is to ensure that the desired, relevant actions are readily perceivable». As McGrenere and Ho (2000: 181) observe, while for Gibson the affordance is the action possibility itself, for Norman the affordance is

the action possibility *and* the way this possibility is conveyed or made visible to a perceiver. In other words, he uses the term affordances to refer to what he calls *perceived affordances*.

The new role given to the act of perception emerges also from the definition that Norman gives for affordances in his best known work, *The psychology of everyday things* (1988):

> The term affordance refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used. A chair affords ('is for') support and, therefore, affords sitting. A chair can also be carried (Norman 1988: 9).

First of all, it is evident that Norman's idea of affordances is not built around the general possibilities for acting, but specifically around the possibilities for acting *on objects* («just how the thing could possibly be used»). It is surely true that artefacts in general, and tools in particular, are a special kind of objects created with the purpose of being used, thus they often have a particular design that provides explicit suggestions about their possible handling and utilization. The intended use of manmade objects is also frequently reflected in their names: in Navajo, a chair is *bikáá'dah'asdáhíor*, i.e. "on-it-one-sits" (MacWhinney 1999: 219), and a towel is *bee 'ádít'oodí*, "one-wipes-oneself-with-it" (Steedman 2009: 186). In this regard we may also mention some familiar words usually cited as examples for transparent compounds, such as *corkscrew* (used to screw corks) or *dishwasher* (used to wash dishes). While Norman restricts and focusses his research on this specific kind of possible actions, related to the usability of objects, Gibson maintains a wider scope, considering that also humans afford behaviours to other humans, animals to other animals, even asserting that also dangerous situations may afford risks to living beings. In this sense, Gibson's idea of affordances recalls a passage in Koffka (1935: 7, quoted in Gibson 1979: 138): «To primitive man each thing says what it is and what he ought to do with it: a fruit says, "Eat me"; water says, "Drink me"; thunder says, "Fear me", and woman says, "Love me"».

This brings us to highlight another major difference between Norman's and Gibson's approaches: for Gibson, affordances are *all* the possibilities for action latent in the environment; instead, for Norman, (perceived) affordances are *only* those objectual properties that may support an activity that are *likely* to be perceived by an agent, according to his beliefs, his goals, and his past experiences (remember the «perceived *and* actual properties of the thing» cited above; emphasis added). Thus, for an agent to perceive an

affordance, it is not sufficient having the physical ability to do this. In a sense, this reformulation (actually a restriction) of the concept stresses even more the idea of complementarity already present in Gibson's work, but also implies a deep difference between the two authors. Norman believes that «affordances result from mental interpretation of things, based on our past knowledge and experience applied to our perception of the things about us» (Norman 1988: 219). It is clear that this assertion contrasts with the idea that affordances exist independently from the perceivers, as stated in the following quotation:

> Despite the interpretations of some psychologists, affordances do not arise as a consequence of mental operations. They are action-referential properties of the environment that may or may not be perceived (Michaels 2003: 137, about Gibson).

While for Gibson affordances point to the *physical abilities* (or effectivities) of an agent, for Norman they point to agents' mental and *perceptual ability*, to their cultural knowledge: the ability to perceive affordances depend also on subjective factors.

But even if Norman shed a new light on the act of perception in detecting affordances, Gaver (1991) goes a step further, separating affordances, intended as possible ways of interacting with objects, and perceptual information about them (Fig. 1.1). Therefore, he recovers the Gibsonian idea that affordances do not need to be perceived, in order to exist.

| Perceptual Information | | Affordance: No | Affordance: Yes |
|---|---|---|---|
| | Yes | *False Affordance* | *Perceptible Affordance* |
| | No | *Correct Rejection* | *Hidden Affordance* |

Figure 1.1. Gaver's classification of affordances. Affordances are separated from the perceptual information that specifies them (adapted from Gaver 1991: 80).

The author also makes an important distinction, dividing affordances into three categories: perceptible affordances (if perceptual information provided by a given object, i.e. its design, matches with the intended use of the same object); false affordances (if an object's design suggests an action that is not the intended one); and hidden affordances (if no perceptual information is provided by the object). Finally, people will usually not think of a given action

if there is no affordance for it, nor any perceptual information suggesting it (correct rejection). The first two cases cover Norman's concept of perceived affordances (McGrenere and Ho 2000: 183).

### 1.2.2 Robotics

In last years, the notion of affordances as first elaborated by Gibson in the field of ecological psychology was also applied to robotics and AI. The idea that an artificial (just as a living) agent can extract information relevant for action directly from the world surrounding it, minimizing the need for complex internal representations, has led researchers to program agents that are more flexible and better able to adapt their behaviour to real world conditions, i.e. embodied agents able to operate in a complex and unstable environment (Horton et al. 2012: 70).

A number of works describe how affordances may be used at different levels of robot control, ranging from perceptual learning to planning. They commonly share a view on affordances as internal relations between external objects and the agent's own actions. For example, Şahin et al. (2007; cf. also Ugur et al. 2009) define affordances as relations that pertain to the robot-environment interaction, and represent them as triples of (1) the agent's behaviour, (2) the object perceived, and (3) the resulting change of state after the agent's behaviour has been applied. In particular, they formalise affordances as interactions of the type (effect, (entity, behaviour)), i.e. an action (behaviour) performed on an entity that produces a given effect. «For instance, the lift-ability affordance is represented as a relation between the (properties of an) object, the behavioural capabilities of the robot and the effects produced by the lift behavior» (Ugur et al. 2009: 178). This formalisation enables artificial agents to record the effects of their actions on perceived objects (i.e. to learn affordances) and to predict a desired effect (from a known behaviour applied on a known entity). This also allows agents to develop planning abilities.

Results obtained in robotics are a good example of how a multifaceted concept like the one elaborated by Gibson can be applied also to other research fields and helps in achieving better results.

## 1.3 Where are affordances grounded? Evidence from behavioural and neurophysiological studies

Despite[2] its various shades of meaning, the theory of affordances emphasises the close correlation that exists between action and perception: it is evident that our possibilities for acting are strongly dependent on our ability to perceive. We couldn't satisfy many of our needs, nor reach most of our purposes (in sum, adapt the world to our necessities), if we were not able to perceive the properties of the environment and the objects surrounding us; but the strong interrelation between action and perception goes beyond our obvious intuition. In the last decades, a number of neurophysiological studies have demonstrated that the motor and perception systems are not isolated, but interact with one another in different ways; moreover, the idea that the perception and sensory-motor systems are also the ground where cognition is rooted (thus, *embodied* cognition; cf. §1.4), is now widely shared among researchers from different fields, so that, nowadays, the embodied approach can act as a unifying element among disciplines (Glenberg 2010). In particular, the close connection between an agent's ability to perceive an object's properties and possibilities for action has been revealed by the discovery of canonical neurons circuits.

Mirror and canonical neurons were first found in the macaques monkeys in the ventral premotor area F5, but later researches provided evidence for the existence of equivalent circuits also in humans' brain (Fadiga and Craighero 2003). Both types of neurons have motor properties and fire when the agent executes a specific action on objects, such as manipulation or grasping. However, they have different visual responses.

Mirror neurons also discharge when agents observe and recognise the same action performed by other agents, or when they hear the related sound; therefore, they are usually intended to play a role in recognition of the others' actions and intentions (Di Pellegrino et al 1992; Rizzolatti et al. 1996; Rizzolatti and Arbib 1998; Rizzolatti and Craighero 2004; Rizzolatti and Sinigaglia 2006).

Canonical neurons were first found in the area F5 of macaques' brain (as mirror neurons), but they fire during goal-directed actions as well as when the monkey simply looks at an object related to action (Rizzolatti et al. 1988; Jeannerod et al. 1995; Murata et al. 1997; Raos

[2] Part of the content of this paragraph has been published in De Felice (2014b).

et al. 2006). Regarding humans, Grafton et al. (1997) conducted a positron emission tomography (PET) and found that the observation of manipulable objects activated the motor system, in particular the left premotor cortex (cf. also Grèzes and Decety 2002). Canonical neurons in the human brain were studied also through neuroimaging techniques. For example, the analysis of Grèzes, Armony et al. (2003) reveals, thanks to the use of fMRI (functional magnetic resonance imaging), the activation of the parietal and premotor areas when subjects passively observe objects, as well as when they execute movements directed to objects. These regions seem to correspond to the circuit in the macaque brain where canonical neurons were discovered[3] (Grèzes, Armony et al. 2003: 933; see also Chao and Martin 2000; other studies will be cited in what follows).

These works demonstrate that perceiving some properties of manipulable objects activates a sort of action simulation in brain circuits. The recruitment of the motor system during object observation (and particularly of the same areas activated during object actual manipulation) has recently been seen as an evidence for the physical and concrete existence of affordances, intended as possibilities for action triggered by object visually perceivable properties.

In particular, Tucker and Ellis (1998) explicitly define affordances, traditionally viewed as possibilities for action, as the motor representations triggered by objects' perception:

> We use the term *affordance* to refer to the motor patterns whose representation visual objects and their properties give rise to, both during explicit goal-directed acts [...] as well as, we argue, before explicit intentions have been formed. Although this is a representational account of affordances, and therefore very different from the use of the term in the ecological sense, it nonetheless has its basis in a similar emphasis of the intimate link between perception and action (Tucker and Ellis 1998: 833).

The authors define their approach a representational one, in that they assume that a mental representation of a visual object also includes encoding of the actions relevant for that object, so that a relationship between the world and objects' representations is

[3] In the same study, an equivalent for mirror neurons circuits was also found, in the concomitant activation of some regions (dorsal premotor cortex, the intraparietal sulcus, the right parietal operculum and the superior temporal sulcus) when subjects observed grasping actions and when they had to imitate them (Grèzes, Armony et al. 2003: 933).

established (Ellis and Tucker 2000: 452). If in Gibson's view (as also in many subsequent works) affordances appear to be anchored in both environment's and agents' dispositional properties, from a representational perspective they appear as «dispositional properties of a viewer's nervous system» (Ellis and Tucker 2000: 466).

The same authors also introduced a different term to refer to the specific motor patterns evoked by object perception:

> The facilitated actions observed in our experiments are of specific components of grasping. Moreover they involve facilitation of particular values of the components concerned. It is not grasping in general that is facilitated, but a specific grasp appropriate to the viewed object. It is a particular shape of the hand and a particular orientation of the wrist, which are afforded. We term these effects, for obvious reasons, micro-affordances (Ellis and Tucker 2000: 467).

This gives us a clear idea of the difference (not only at a terminological level) from the Gibsonian tradition. Here, "higher level" affordances, intended as actions associated with an object's function, are clearly distinguished from "low-level" affordances, that refer to the minimal motor patterns potentiated by objects, such as the specific components of grasping evoked by an object with particular features (see also Tucker and Ellis 2001).

This definition of affordances takes into account both the perceiver's motor capacities (also because the brain motor patterns are dependent on agents' actual possibility to act, as will be shown in next paragraphs)[4] and the visual information available (objects' properties, such as dimension, shape, location etc.).

Therefore, on the one side, psychologists and cognitivists reshaped the concept of affordances taking into account the recent discoveries in the neurosciences, as well as, on the other side, neuroscientific research often explicitly quotes affordances and refers to psychological works. In particular, object manipulation and object grasping is nowadays the field in which the overlapping of different theories and methods related to affordances is most evident. On this specific topic, many efforts from different disciplines (also neurolinguistics: cf. in particular §1.4.2) converge. The variety that characterises the huge amount of research activity conducted on the specific action of grasping objects well reflects

[4] In this regard, the more general mental faculties of an agent have also to be taken in account: most of the studies cited so far are conducted on healthy subjects.

the need for a multidisciplinary approach for analysing the complex relation between action and perception.

The remaining part of the chapter will focus on some phenomena emerged from behavioural and neurophysiological research that assume great relevance in light of the purposes of this thesis. Given the high number of studies related to the grasping worth to be mentioned, the literature review will be organised around few topics, each one devoted to a single feature that has been proved having a role in modulating brain activity, affecting the motor representations the object raises and thus being responsible for recruiting motor responses in a selective way. This will give us the chance to mention the researches that have a major impact on the theory of affordances.

### 1.3.1 Object's dimension and shape

The type of grip afforded by an object, according to its shape and dimension, is able to influence motor responses.

Usually, in behavioural research, affordances related to a particular objectual feature are studied by verifying if this feature has an impact on a task, for which the same feature is not relevant. In Tucker and Ellis (2001), participants were asked to perform a categorization task, i.e. to choose whether an observed object was an artefact or a natural kind. They had to express their judgement performing a unimanual precision or power grasp on a manipulandum. Most importantly, the requested grasp could be congruent or incongruent with the object's dimensions: there were small stimuli (e.g. a screw, a hammer), affording a precision grip, as well as large stimuli (e.g. a grape, a cucumber), affording a power grip. Measurements of reaction times indicated that motor responses were significantly affected by the compatibility between the type of grasp requested by the task and the type of grasp afforded by the visually presented object: when participants were viewing large objects, power responses were faster than precision responses, whereas for small objects, precision responses were faster than power responses.

Similarly, in 2002, Gentilucci demonstrated with a behavioural study that some object-intrinsic properties, such as the volume (in particular the size of an object's graspable part) and the shape, affect grasp kinematics, even when these components are irrelevant to the task. In particular, the author shows (Gentilucci 2002: 1150-1151) that subjects grasped two bells of different volume (the larger one was approximately of 7 cm, whereas the smaller

one was approximately 3 cm) in different ways, even if their stalk, that was the part where subjects were explicitly requested to grasp the object, was identical, i.e. had the same physical features. Although people were asked to grasp the stalks, the volume of the two objects influenced the grasp kinematics: the hand, when the target-object was the large bell, was partly preshaped as if the entire object, and not only its stalk, was to be grasped. In particular, the volume effect was significant in two of the three parameters considered for the grasp: peak velocity of finger aperture and maximal finger aperture, but not for the percentage of time to maximal finger aperture. These results support the hypothesis that a single object's motor representation, which codes *all affordances* enabled by the object, is involved in grasp kinematic implementation: in this experiment, kinematics turned out to be influenced by a task-irrelevant characteristic such as the volume of the object, even if subjects perfectly knew that only the stalk had to be grasped.[5] Other details from Gentilucci's experiments will be given at §1.3.4.

These data point to the fact that objects' shape and dimension affect motor representation of objects (affordances) that are responsible for hand shaping and grasp kinematics during the execution of a grasp (see also Girardi et al. 2010):

> Intrinsic object properties influence both the selection of the type of grip and the grasp kinematic implementation. These properties are referred as to object affordances, i.e. motor representation eliciting particular types of interaction with the object (Gentilucci 2002: 1139).

Other interesting results about the role of object physical shaping and volume come from brain imaging studies. Grèzes, Tucker et al. (2003) conducted a complex behavioural and fMRI experiment in which participants were asked to execute a power or a precision grip

[5] The author assumes that the AIP (anterior part of intraparietal sulcus, that contains motor-dominant neurons, active during the entire motor act of grasp, as well as visual-dominant neurons, that fire during object observation, sometimes reflecting different object's shape and size) sends visual signals of object properties to area F5, active during specific phases of grasp and selective for a particular type of grip. Area F5 receives information about object properties, selects the type of grip and hand movements, and sends back a motor command to the area AIP. This loop would be responsible for the matching of the hand grip with the features of the target object and its affordances. The author suggests, in the light of his experimental results, that the AIP-F5 circuit extracts not each affordance separately, but concurrently all the possible affordances from a presented object (Gentilucci 2002: 1152-1153).

response on a manipulandum, according to the type of object they were presented, either natural or manmade. The study aimed to find a compatibility effect between the type of grip afforded by the object, that was a task-irrelevant feature, and the type of grip requested as response by the classification task. The type of grip afforded by stimuli objects, in the experiment, was exclusively related to their dimension: authors used large objects to afford a power grip, and small objects to afford a precision grip (cf. Tucker and Ellis 2001). As expected, responses were faster in congruent cases, i.e. when the type of grip afforded coincided with the grip that participants had to perform (a power grip response to a large object, or a precision grip response to a small object). Brain imaging techniques also revealed that the degree of motor activation (recorded in the parietal, dorsal premotor and inferior frontal cortex) during the execution of a given hand grip depended on the congruence between the hand grip afforded by the object and the grip requested by the task. During incongruent tasks, a strong competition between the action requested (e.g. precision grip for artefacts categorization) and the action afforded by the stimulus object (e.g. power grip for a large artefact, as a hammer) is generated, and this competition is responsible for both the slower reaction times in motor responses[6] and the greater activation of sensory-motor system. The study provides evidence for the fact that different object dimensions may automatically generate different types of motor responses and, in general, that object affordances are evoked even when they are irrelevant to the task: in this experiment, as in those previously mentioned, no reaching and grasping movement towards the object was actually required.

### 1.3.2 Affording parts

We will use the expression *affording part* to refer to the part of an object that more than others is typically involved in actions (e.g. a handle) and elicit affordances (see *infra*). Behavioural studies, mostly based on compatibility paradigms, provide evidence for the effect of spatial alignment on affordances activation: a task-irrelevant objectual property, such as the orientation of its affording part, may potentiate the execution of hand motor acts, and

[6] Authors are not able to say whether the difference in reaction times is due to a facilitation effect occurring in congruent trials, an interference effect occurring in incongruent trials, or both (Grèzes, Tucker et al. 2003: 2738).

this happens when the orientation of the affording part is spatially aligned with the responding hand (compatibility effect).

For instance, in Tucker and Ellis' (1998: 834-838) experiments, the leftward or rightward orientation of familiar objects (such as a teapot, a frying pan, a jug, a kettle) had a significant effect on the speed (measured as reaction time) with which a left or right response was executed by, respectively, the left or the right hand. In the experiments, the horizontal orientation of the object was irrelevant to the task requested: participants just had to decide whether the household item they were seeing in a picture was upright or overturned and respond accordingly with a key press performed with their left or their right hand. Push-button responses were faster and more accurate when they corresponded with the orientation of the object: for instance, objects with a leftward-oriented handle improved the performance for left hand responses. It is worth noting that in a second experiment participants were asked to execute left/right responses by using not two hands, but only two adjacent fingers of their right hand. In this case, no compatibility effect occurred, demonstrating that «it is the affordance for grasping *by a particular hand* that gives rise to the binary left-right distinction» (*ibid.*: 838, emphasis added), not the object orientation itself.

Similar effects also emerged from other compatibility studies, such as Riddoch et al. (1998), Ellis and Tucker (2000), Tucker and Ellis (2001), Phillips and Ward (2002). What all these studies have in common is the fact that they use stimuli that have an obvious action connotation (they are all concrete, familiar and meaningful objects) and are asymmetrical (thus visually salient areas might bias attention), such as handled cups. In order to avoid confusion with semantic or attentional factors and to investigate if orientation can be considered a "pure physical affordance", i.e. «an affordance that is solely revealed by the physical structure or arrangement of the object» (Symes et al. 2007: 239), Symes et al. (2007) studied the orientation-dependent compatibility effect adopting a series of elonged, geometrical stimuli. Such stimuli could be oriented $\pm 45^{\circ}$ from the perpendicular (i.e. with a left-down or right-down orientation) and consisted of an abstract 2D rectangle, a 3D cylinder oriented only in the frontal plane, a 3D cylinder also rotated in depth (in order to appear to be pointing out in space towards a particular hand of the viewer). Orientation was again a task irrelevant property, since participants had to press a right or left button if the stimulus pattern was straight or wobbly, respectively; if the pattern was neuter, they had to wait until it changed into a wobbly or straight pattern. Reaction times from motor responses in congruent (i.e. when the responding hand corresponds to the orientation of the object) or

incongruent trials were recorded. From the results obtained in five different experiments conducted on these stimuli, the authors conclude that the more realistic, three-dimensional and graspable an object appears to be, the more potent its pure physical affordance is. The 3D cylinder oriented only in the frontal plane produced a small orientation-dependent spatial compatibility effect, whereas the abstract 2D rectangle that had the same orientation did not. For the 3D cylinder rotated in depth, pointing out in space towards a particular hand of the viewer, larger and more robust compatibility effects were produced (*ibid.*: 251-252).

These findings from behavioural studies may be compared with the brain imaging analysis conducted by Buccino et al. (2009). The authors investigated, using single pulse TMS, the activation of parieto-premotor circuits while subjects observed photographs of familiar handled objects (six common containers, such as a mug, a coffee maker). The handle of these objects-stimuli could be rightward- or leftward-oriented; crucially, it could also been broken or intact. In their study, the largest motor evoked potential area was recorded from hand muscles when participants (that were all right-handed) looked at images of rightward-oriented objects provided with intact handles. When the handle was leftward-oriented, or when it was rightward-oriented but broken, motor activation was much less evident.

### 1.3.3 Spatial constraints

Many studies show the recruitment of the motor system during the observation of graspable objects. But it is evident that actual manipulation of objects may be realised only if objects are close enough to agents so that they are able to reach them. Thus affordances, intended as possibilities for action, should depend not only on the intrinsic relation between objectual features and agents' abilities, but also on the real possibility for the agents to act on the objects and so, first of all, to reach them with their hands.

It is known that the way we perceive the space surrounding us is highly dependent on subjective or contextual factors. For instance, the judgement of the distance with respect to an object varies according to the action abilities of the agent; the distance is judged to be shorter when the agent has the possibility (as well as the intention) to reach and grasp the object with a tool (Witt et al. 2005), or when handle orientation makes the object easier to be picked up (Linkenauger et al. 2009). Thus, perception of a fixed distance between an agent and an object is modulated by the dispositions of both the object (e.g. orientation) and the subject (e.g. capabilities, intentions).

Interestingly, in some works, it is the distance between the object and the agent that is modulated by researchers, in order to investigate the possible existence of spatial constraints on affordances. For instance, Costantini et al. (2010), adopting a spatial alignment effect paradigm, conducted a study in which participants had to replicate a reach and grasp movement as soon as they saw a task-irrelevant go-signal, i.e. a mug on a table. The handle of the mug could have a congruent or incongruent orientation with respect to the action required, and the object could appear in the reachable (30 cm) or unreachable (150 cm) space, with respect to the agent. Thus, as the stimulus appeared, two different motor patterns were recruited: the representation of the grasping movement to execute and that of the grip afforded by the object. The analysis of the grasping onset time showed a compatibility effect (shorter reaction time) only when the object fell within the reachable space.

Other experiments confirm with neural evidence what is noticed at a behavioural level. Cardellicchio et al. (2011) investigated exactly whether motor representations depend not only on the visual presentation of the affording feature of an object, but also on its reachability, i.e. on the distance of the affording object with respect to the observer. In their TMS experiments, subjects observed a 3D room with, again, a mug on a table, or a big box, either located in the reachable space, or outside the reachable space; their left primary motor cortex was stimulated and motor evoked potentials (MEPs) were recorded for the right hand (right first dorsal interosseus and opponens pollicis). MEPs were higher in amplitude when the mug was presented in the subjects' peripersonal space, and not in the extrapersonal space. Notably, when the object was non-graspable (a large box), a similar effect was not observed.

It is worth noting that no significant difference in measurements was registered with relation to different (left/right) orientation of the handle of the mug, so results for these two conditions were merged. This contrasts with results from Buccino et al. (2009) where a significant difference was found between the MEPs registered in the right intact handle condition compared to those registered in the left intact handle condition. This is explained by the authors (Cardellicchio et al. 2011: 1371) considering that the experiment conducted by Buccino et al. (2009) clearly focussed on the orientation on the handle and that this could have attracted the attention of participants on this specific structural feature.

These studies reveal that the processing of an object's affording features is spatially constrained, i.e. it depends on the spatial relationship between the object physical features and the individual's motor abilities. In other words, the relation of affordance:

> […] depends on a further relation between its *relata*, that is, a spatial relation which is not constitutive of the distinctiveness of the affordance but makes it possible. In order for something to be graspable with respect to an individual endowed with the appropriate motor abilities, it has to fall within his or her own peripersonal space – better, it has to be ready to her own hand (Costantini and Sinigaglia 2012: 440).

Finally, other two works are worth mention. In Costantini et al. (2011) and Cardellicchio et al. (2013), much of the experimental setting and procedure already described was reused (the mug on a table in the peripersonal or extrapersonal space), but an avatar was introduced in the scene. Interestingly, the spatial alignment effect, as well as the highest motor-evoked potentials from subject's hand muscles, were observed both when the mug was either reachable for the participants, and when it was unreachable for them but reachable for the avatar. In this condition, the affordance relation is mediated by the peripersonal space of another individual. According to the authors, this effect is due to the existence of a mirror mechanism[7] that maps the peripersonal space of others onto the observer's own peripersonal space, as well as others' action potentialities onto the observer's own motor abilities. The mirror mechanism for the peripersonal space also shed a new light on social cognition (Costantini and Sinigaglia 2012: 452-453):

> The fact that the affordance relation is not a private business of a single individual, but it relies on a mirror mechanism that allows one to share the space of her own action with others, highlights that the investigation of affordance mandatorily involves dealing with the cognitive processes underlying basic social cognition. […] The space mirror mechanism provides us with an immediate precomprehension of their own body as an acting body as well as the effective range of their bodily agency. Such precomprehension appears to play a critical role in action understanding, at least at the basic level of the motor-based action and intention understanding, thus allowing to highlight the very first steps in our making sense of others as well as in our sharing a common world with them.

[7] As already stated, mirror neurons fire when an agent executes a specific actions on objects, such as manipulation or grasping, as well as when the agent observes (or hears) the same action being performed by another agent; for this reason, they are usually intended to play a role in recognition of the others' actions and intentions (Rizzolatti and Craighero 2004).

### 1.3.4 Beyond the object: familiarity

As Buccino et al. (2009) point out, familiarity with objects and with the actions they are typically involved in is an important factor that influences motor representations (Buccino et al. 2009: 3077). This has also been highlighted, from a slightly different point of view, by Gentilucci (2002: 1152):

> Familiar objects automatically activate habitual types of interactions, which very strongly influence grasp implementation. It can be argued that, from a motor point of view, familiar objects can be represented by the type/types of interaction that we habitually have with them.

In his behavioural study, the author conducted eight experiments in which participants had to reach and grasp various objects, different for weight, volume, shape, intrinsic height and centre of mass. The analysis indicates that familiarity affects the grasp kinematics, because the volume effect (cf. §1.3.1) is stronger when subjects have to grasp familiar objects (e.g. fruit, bell) than unfamiliar ones (geometrical solids of different shapes). For instance, we can consider two of the experiments described in Gentilucci (2002), in particular Experiment 1, involving an apple and a strawberry, and Experiment 4, involving two spheres of different dimensions. The apple and the big sphere were green, had the same weight and a diameter of around 7.0 cm, while the strawberry and the little sphere were red, had the same weight and a diameter of around 3.0 cm. All objects had an identical stalk of around 2.0 cm height that participants were requested to grasp, so all conditions but the type of object (fruits vs. geometric solids) were equivalent. The volume effect (already discussed in §1.3.1) on hand shaping during the reach and grasp movement of small vs. large objects turned out to be significant, both for the two fruits and for the two spheres: objects' dimension affected hand shaping and grasp kinematics during the execution of the grasp, even if in both experiments the object's part that subjects had to grasp had the same physical features. However, what is most important here is that the volume effect was much more relevant for fruits than for spheres; thus, *ceteris paribus*, according to authors this difference in results is due to objects' familiarity.

### 1.3.5 Stable and variable affordances

The possible properties of an object mentioned so far, namely its spatial location with respect to an agent, the familiarity that an agent has with it, its dimension and shape, or the presence of an affording part, are all characteristics that are able to modulate the sensory-motor system, as the neurophysiological and behavioural researches reviewed so far demonstrated, but they are not equivalent to each other. We may take as an example a very common stimulus, such as a cup. If a perceived cup has a handle, it will activate stronger motor simulation in the brain, with respect to a non-handled cup, or to a cap with a broken handle (as demonstrated, for instance, by Buccino et al. 2009; cf. §1.3.2); the presence of what we have called an affording part is a physical and invariant property of the object. Then, we also know that, if this handled cup is presented to a right-handed subject, it will activate, again, stronger motor simulation if it is rightward-oriented rather than leftward-oriented; but this characteristic to be a rightward-oriented object is less stable (cf. *infra*), in comparison with the presence of a structural part designed for grasping. It is not an object's intrinsic characteristic. Rather, it depends on the situational context and emerges in relation to the agent: therefore, orientation is a property of the object only in relation to a perceiver.

Borghi and Riggio (2009), on the basis of empirical evidence, propose to distinguish *stable affordances*, that are usually linked to invariant features or properties of objects and incorporated into an object's representation, from *variable affordances*, that are related to temporary object characteristics and are specific of a given situation. When a temporary property, such as orientation, is strongly associated with the typical actions we perform on the object (right-handed people usually take cups by their handle, if it is rightward-oriented), the property is said to be a *canonical affordance*.

## 1.4 Affordances and the embodied language

We share with many of the authors cited so far the idea that affordances can be better understood if defined as the motor patterns automatically triggered by objects' perception (cf. §1.3; further evidence will be provided in this paragraph). This allows us to ground affordances in the perception-action system and to benefit from a number of neuroscientific findings that also shed new lights onto human cognition and behaviour (for instance, consider the impact that the discovery of mirror neurons had on social studies). We embrace

this definition, which stems fundamentally from an embodied approach to the concept of affordances.

Embodiment theories share the idea that our cognitive structures originate from our concrete experience with the world surrounding us. Since experiencing the world means, for us, first of all to perceive it with our sensory system through different modalities, and then also to concretely interact with it (moving ourselves in the surrounding space, acting on objects, and so on), cognitive structures develop from perception and action (Pecher and Zwaan 2005). An increasing body of works (e.g. Barsalou 1999; Gallese and Lakoff 2005) has recently indicated that language, at least to some extent, may also be embodied. From this point of view, words and sentences are not seen as abstract, amodal and arbitrary mental symbols, as assumed by symbolist theories (e.g. Fodor 1975; Pylyshyn 1984; Mahon and Caramazza 2008; Chatterjee 2010), but they are grounded in the real world and in human experience: linguistic concepts consist of mental "simulation" of the experiences which the words and sentences refer to, and linguistic material is thus processed with the same brain mechanisms that underlie perception, action, emotion and other types of human-world interactions. These two approaches are not necessarily mutually exclusive: for example, Paivio's (1971) dual-coding theory assumes that concrete words are processed by both the symbolic and the embodiment systems, while abstract words are processed by the symbolic system only. Other researchers, on the contrary, consider that both abstract and concrete language could be grounded in situated knowledge (e.g. Barsalou and Wiemer-Hastings 2005).

In what follows, we will comment upon the concept of embodied language, presenting some of the works conducted in this field that support this view and are in line with the studies mentioned in the last paragraphs.

### 1.4.1 Action-related verbs and sentences

The fact that language processing passes through motor simulations, at least in the case of language related to action, appears evident from a growing body of both behavioural and neuroscientific research conducted on sentences or verbs related to actions (e.g. Barsalou 1999; Gallese and Lakoff 2005; Pulvermüller 2001; 2002; 2005; Pulvermüller et al. 2005; Tettamanti et al. 2008; Jirak et al. 2010). The main finding of these works is that reading or listening to action-related verbs or sentences activates the same motor and premotor brain areas that are activated when subjects perform the actions denoted by the verbs or sentences

considered (for a review, see Pulvermüller et al. 2009): we may refer to this phenomenon as to the meaning-action matching (De Vega 2012).

A clear and direct evidence for the meaning-action matching is provided by Hauk et al. (2004): in this study participants passively read action-related verbs denoting mouth, hand or leg movements (e.g., to *lick*, *pick*, or *kick*); in a different task, they also performed actions involving the same body parts (moving tongue, index fingers and feet). Using fMRI, the authors found that these action verbs, pertaining to different semantic subcategories, activate the motor cortex in a somatotopic way (i.e. activate specific regions of the cortex responsible for the motor control of different areas of the body). Language-related cortical activity overlaps with the diverse activation patterns observed in premotor and motor cortex during actual movements of the body parts words refer to (Hauk et al. 2004: 301).

The involvement of the sensory-motor system in language processing is also proved in an indirect way at a behavioural level, by the facilitation or interference effects on motor responses caused by the comprehension of action-related sentences (the action-sentence compatibility effect). For instance, the meaning of actional sentences may facilitate subsequent actions that require performing movements congruent with the type of action described by the sentence. Thus, a movement of the hand towards (or away from) the subject's body will be performed faster after understanding sentences describing movements directed towards (or away from) the subject, such as *Courtney handed you the notebook* vs. *You handed Courtney the notebook* (Glenberg and Kashak 2002; see also Borghi et al. 2004; Zwaan and Taylor 2006).

However, as we said, the meaning-action matching may also cause interference (rather than facilitation) effects. Buccino et al. (2005) used transcranial magnetic stimulation and a behavioural paradigm to assess whether listening to action-related sentences modulates the activity of the motor system. The authors found that, after hearing foot-related sentences, subjects' motor responses were faster when performed by hands than when performed by feet (and the amplitude of motor-evoked potentials recorded from subjects' foot muscles decreased). Conversely, after hand-related sentences, their foot responses were faster (and the amplitude of motor-evoked potentials recorded from their hand was reduced). This interference effect, reflected in slower motor responses, reflects a neural competition due to the fact that the same effector (the hand or the foot) is involved in both the motor response and the meaning of the sentence.

Similar facilitation or interference effects also emerged for emotion-related sentences, the processing of which interacts with the muscles for facial expressions (Havas et al. 2007; 2010).

Despite growing evidence that cortical motor areas are activated during the comprehension of action-related verbs and sentences, the type of relation that links the two phenomena still remains object of debate, whether they are just co-occurring (for different possible reasons), or the recruitment of the motor-system really contributes information that is critical in order to comprehend sentences (on this, see also the discussion in Mahon and Caramazza 2008; De Vega 2012). There are at least two reasons that may favour the latter hypothesis. The first one, is the precocity of the motor representations. As a matter of fact, as Pulvermüller et al. (2009: 87) put it, neurophysiological researches «confirm near-simultaneous early brain correlates of phonological, lexical and semantic information immanent to a spoken word within the first ~150 ms [milliseconds] after the auditory input allows for word identification». According to Buccino and Mezzadri (2013), a very early and likely automatic recruitment of the motor system during language processing (150-170 ms after the stimulus, as many of the studies cited so far record) provides a compelling evidence for the necessary and crucial role of the motor system in language comprehension (as it appears from interference effects). On the contrary, other kinds of phenomena that occur later (such as facilitation effects) could be due just to an interaction between the motor system and the language, but not as a necessary part of language processing.

The second piece of evidence comes from studies conducted on patients with Parkinson disease: people that have damage in their motor area, also have selective difficulties in understanding verbs (Boulenger et al. 2008; Castner et al. 2008; Crescentini et al. 2008). Furthermore, a recent study conducted on healthy subjects interestingly shows that a reversible disruption of the premotor cortex inducted by repetitive TMS, that causes a sort of "virtual lesion", interferes with the comprehension of sentences describing manual actions (Tremblay et al. 2012). These results confirm that cortical motor regions are critical to word understanding and that processing lexico-semantic information about action words necessarily depends on the integrity of the motor system (Boulenger et al. 2008: 743).

As many researchers suppose, the partial overlapping of the motor patterns activated during the comprehension of action-related verbs and sentences and those activated during action execution may be due to an involvement of the mirror neurons system in the processing of action-related language (for a review, see Kemmerer 2006; cf. also Buccino and Mezzadri 2013). The mirror neuron system is organised in a somatotopic fashion (Buccino

et al. 2001; Wheaton et al. 2004), as somatotopically differentiated are the motor simulations triggered by linguistic stimuli as well (i.e. stimuli activate specific regions of the cortex responsible for the motor control of different areas of the body). A clear evidence for the involvement of the mirror neurons system in the processing of action-related language is provided, for instance, by Aziz-Zadeh et al. (2006), that found that the same brain areas are activated both when subjects read sentences related to actions, and when subjects observe those actions being performed by other agents.

### 1.4.2 Action-related nouns

If the circuit of mirror neurons is probably involved in the comprehension of action-related sentences and verbs, it is also possible that canonical neurons [8] underlie the comprehension of action-related nouns. As a matter of fact, the visual perception of graspable objects activates the same neural system activated during their actual manipulation (§1.3). Is a similar recruitment of the motor region also observed when the stimulus is a word, denoting an object, and not an image or an object? Neurophysiological researches demonstrated that nouns and verbs activate different neural circuits (cf. *infra*; for a review, see Vigliocco et al. 2011) and that the same is true also for some subcategories within the class of nouns. These findings strongly suggest that linguistic representations, as well as the concepts associated to them, give rise to neural patterns that are heterogeneously distributed and integrated in different cortical areas.[9]

Unfortunately, there are not many studies specifically conducted on the linguistic category of nouns, especially at a neurophysiological level. But we know, for instance, that brain regions that become active during processing of concrete nouns are not the same as those activated by abstract nouns (e.g. Kiehl et al. 1999; Martín-Loeches et al. 2001); furthermore, the conceptual knowledge about concrete nouns of different semantic categories, such as

[8] We should recall that both mirror and canonical neurons have motor properties, i.e. fire when an agent executes a specific action on objects, such as manipulation or grasping, but they have different visual responses. Mirror neurons also discharge when the agent observes the same action being performed by another agent, whereas canonical neurons also fire when the agent simply looks at an object related to action (cf. Grèzes and Decety 2002; Grèzes, Tucker et al. 2003; Chao and Martin 2000).

[9] A similar phenomenon, for example, regards the representations of faces and objects of various categories in the ventral temporal cortex, that are widely distributed and overlapping (Haxby et al. 2001).

animals, fruits and vegetables, tools, appears distributed in different neuroanatomical areas. In particular, the action of naming graspable artefacts (especially tools) involves the cortical areas where canonical neurons were found in monkeys and humans, the same areas that are also activated by objects manipulation (e.g. Grabowski et al. 1998; Chao et al. 1999; Caramazza and Mahon 2003).

There is an increasing body of research demonstrating that the recruitment of the motor system is particularly evident during the presentation of noun words referred to graspable artefacts. For instance, in a positron emission tomography (PET) study conducted by Martin et al. (1996), the left ventral premotor area turned out to be selectively recruited when subjects named pictures of tools, whereas names of animals activated a different area, the left medial occipital lobe. Similar results were also obtained with fMRI (Chao and Martin 2000).

Most of the works cited thus far in this paragraph are based on tasks in which subjects are required to name pictures of objects, thus they receive a visual stimulus and have to actively produce a linguistic response, namely retrieving the correct word for each stimulus. Although, some researches indicate that nouns denoting tools or artefacts, as action-related verbs, selectively activate the motor system even without any visual stimulation that involve object's images. For instance, Cattaneo et al. (2010) conducted an interesting behavioural and neurophysiological study that points out the recruitment of the ventral premotor cortex (the probable homologue of macaques' F5 area where canonical neurons are found) in the processing of tool-related words. Subjects were primed by reading either the word "Tool" or "Animal" (or a sequence of symbols, in the control condition) written on a screen, after which a common word denoting an entity of one of these categories appeared. Single-pulse TMS was applied at each target onset over either the left ventral premotor cortex (supposed to interfere with the processing of tool words) or the left dorsal premotor cortex (supposed not to have any effect on tool words processing). Participants were required to decide as fast as possible whether the stimulus-word shown after the prime denoted a tool word or an animal, by pressing one button or another with their left hand.

As expected, from a behavioural point of view, a facilitation effect of the prime over responses reaction times was found: subjects responded faster to the target words when they were congruent with the primed category, relative to the condition in which the target words were incongruent with it. Furthermore, TMS results show that when the target word denoted a tool, TMS applied over the ventral premotor cortex interacted with the prime effect,

facilitating reaction times when the prime was "Animal" (because TMS increased the cortical excitability of the ventral premotor cortex responsible for the comprehension of tool words), but having no effect when the prime was "Tool" (the primed neuronal representation was less susceptible to the facilitatory effect of the stimulation). No effects were observed when the target word denoted an animal or when TMS was applied over the dorsal premotor cortex. Taken together, these results clearly indicate that the left premotor ventral cortex (the probable homologue of macaques' F5 area where canonical neurons are found) contains a representation of the "Tool" category, but not of the "Animal" category, and is involved in the analysis and comprehension of words denoting graspable artefacts.

The activation of the motor system during the processing of nouns denoting graspable artefacts is also demonstrated by Gough et al. (2012). The authors applied TMS to the primary motor cortex representation of the first dorsal interosseus muscle of the right hand after the subjects read noun words, to study the differences in terms of motor representations between nouns denoting artefacts or natural kinds. Motor evoked potentials were larger for names denoting graspable artefacts (tools), with respect to those evoked by nouns referred to natural objects.

Thus, the modulation of cortical motor regions during noun words presentation is comparable to the activity observed during the visual perception of the corresponding objects or their images. In line with this assumption, Shinkareva et al. (2011) show that words and pictures may give rise to common neural representations. In particular, graspable artefacts are a special class of stimuli that stand out for the motor simulations they evoke.

Behavioural studies conducted on linguistic material support these findings and, from a more general point of view, indicate an involvement of the motor system in the semantic processing of nouns. This activation may clearly emerge from interference effects (cf. §1.4.1). For instance, Glover et al. (2004) conducted a kinematic study in which participants had to read the names of objects of different sizes that afford a power or a precision grip (e.g. *pea*, *grape*, *pencil* vs. *apple*, *orange*, *baseball*) and then, after an acoustic signal (1 second after stimulus presentation), to grasp a wooden block. The authors found an interference of the type of grip afforded by the object denoted by the noun on the grasping movement directed to the block. For example, reading a word representing a large object led to a larger grip aperture than reading a word representing a small object. Affordances evoked by a word are able to influence the planning of grasp kinematics: this finding is quite similar to the results obtained with visual and concrete stimuli (cf. §1.3.1, in particular about the experiments of Gentilucci

2002; see also Bub et al. 2008, in which both depicted objects and visual word are used as stimuli).

Another significant interference effect was found by Marino et al. (2013), that carried out a go/no-go experiment on Italian nouns referring to hand-related (e.g. *forbici* "scissors", *spazzola* "brush", *forchetta*, "fork") or foot-related (e.g. *pedale* "pedal", *pattini* "skate", *scalinata* "staircase") objects and abstract entities (e.g. *superbia* "arrogance", *gelosia* "jealousy"). Participants read nouns on a computer screen and were required to give a motor response, as fast as possible, pressing a key with the left or the right hand, if the noun denoted a concrete object. The go-signal might appear early (150 ms) or late (1150 ms). One of the most interesting results was that responses were slower when participants had to respond with the right hand to hand-related nouns; this interference effect appeared only for the early go-signal and is likely due the early involvement of the motor cortex of the left hemisphere in the representation of artefacts activated by words (cf. Buccino et al. 2005, for a similar effect).

Research conducted on linguistic material thus points out that motor simulations are involved in the comprehension and the processing of action-related verbs, nouns and sentences: it also stands to reason that the concepts of manipulable objects, keeping track of our direct experience, automatically activate affordances, i.e. motor information regarding micro-interactions with their referents (on this topic, cf. Borghi 2005; 2007; for a different opinion see Mahon and Caramazza 2008).

Therefore, as stated by MacWhinney, affordances are thoroughly grounded in both the motor and the sensory systems, and understanding individual words reactivates our normal, personal encounters with objects (1999: 218-219):

> These encounters involve both motoric actions and sensory perceptions. When we hear the word *banana*, we activate neural pathways that are involved in our nonfictive interactions with real bananas. In this sense, understanding the meaning of an object involves running a "cognitive simulation" of our interactions with that object in terms of its most salient affordances.

## 1.5 Concluding remarks

A body of works provides neurophysiological and neuropsychological evidence that the motor system is activated not only by the visual perception of manipulable objects (e.g. Grafton et al. 1997; Binkofski et al. 1999; Chao and Martin 2000; Grèzes, Tucker et al. 2003; Buccino et al. 2009), but also by the nouns denoting them (Cattaneo et al. 2010; Gough et al. 2012). Behavioural studies support these results (Glover et al. 2004; Marino et al. 2013).

Affordances theory finds a fertile soil in these contributions. Motor simulations activated by objects' perception can be intended as affordances (or micro-affordances), in that they function as true «preconditions for action» (Greeno 1994: 340), as demonstrated by the effects of interference or facilitation often recorded in behavioural studies. Motor simulations, emerging as a sort of memory of past experience, not only allow to *understand* a stimulus, but they also prepare actions. It is evident the connection with Gibson's idea of affordances: when we see a graspable object in our peripersonal space, our neural system is immediately prepared to grasp it, rising motor patterns as images of possible actions in a fast, automatic and somatotopic fashion. It is in this mechanism of mental imagery and simulation that possibilities for action are rooted.

The fact that linguistic material is able to modulate the motor system in a similar way means that words, at least noun words referred to graspable artefacts, are embodied; moreover, it also demonstrates that noun words, as well as objects, may automatically trigger affordances.

Nowadays, one of the major contributions to the discovery of the close connection between action and perception, as well as of the embodiment of language, was provided by the research on grasping. Resting upon these considerations, the remaining part of this thesis will focus on this kind of event. The term *affordance* will be used to refer to the sensory-motor patterns activated at a neural level by object perception, that are activated also, through a motor imagery system, by the word denoting the object itself. *Affording properties* will be those specific physical features (whether variable or stable) of objects, such as dimension, shape, constituent parts, location, that are able to trigger and modulate specific motor responses; in particular, the *affording part* will be the part of an object typically involved in (and often specifically designed for) the grasping, which is more likely to affect such brain activity.

# Chapter II

# The Affordance of Grasping: an Experimental Study

This chapter presents an experiment conducted with the main purpose of investigating the linguistic reflexes of the affordance of graspability. The sections in which it is divided describe the purposes of the study (§2.1), the theoretical background (§2.2), the methods adopted to conduct the experiment and to transcribe the oral interviews (§2.3; §2.4). The analysis of the transcripts and the results will be discussed in the next chapters.

## 2.1 Purposes

In the previous chapter, a definition of affordances that considerably relies on neuropsychological findings was outlined. Affordances can be defined as possibilities for action: they are substantiated by the motor information automatically activated by objects' perception and encoded in the object's representation (e.g. Tucker and Ellis 1998; Symes et al. 2007; Buccino et al. 2009; Cardellicchio et al. 2011; Borghi et al. 2012). Since these evoked motoric patterns are able to influence subsequent actions directed towards the object itself,

affordances are true preconditions for activity. Moreover, the physical and situational properties of an object that are able to modulate the sensory-motor system, especially the very same areas that are activated during objects' actual manipulation and grasping, were defined as objects' affording properties.

Despite being all related to grasping, most researches cited in the previous chapter focus on very specific issues, for example measuring the effects of an object's dimension, object's spatial location with respect to the agent's location, object's semantic type, object's shape and constituent parts, and so on. Therefore, it is quite clear that affordances are modulated by a number of different parameters, but it is not easy to consider them all together, in a unitary framework. What makes an object graspable, in the end? It would be rather difficult to define what the affordance of graspability is, whereas it is easier to identify and describe the specific affording properties related to the grasping.

Operationally, in this thesis, the affordance of graspability is defined as extensionally corresponding to the set of the (affording) properties of the object that make it graspable. It is also assumed that perhaps not all the affording properties of graspable objects that were mentioned in the previous chapter (e.g. spatial relations, dimension, and presence of affording parts), from a more psychological perspective, have the same status in the awareness of agents and are perceived as contributing in the same way to objects' graspability.

This research originates from the conviction that language may open a window into this interesting issue. Therefore, the first question this thesis aims to answer to is: does language reflect affordances? In other words: is language sensitive to the same variables that modulate other aspects of human behaviour?

This chapter describes an experiment primarily designed to explore how the reference to the graspability of an object can be made. Crucially, this information is derived from linguistic descriptions of grasps (such action description task will be illustrated in detail in §2.4). This is, to our knowledge, a hitherto unattempted task that may reveal a new way to investigate affordances, which are rarely mentioned in linguistic studies. The next paragraphs will present the theoretical background and the methodological framework that guided the experiment design.

## 2.2 Theoretical and methodological issues

The purpose of the study has been just introduced, but every constituent part of the experiment also has to take into account, and rely on, theoretical bases. Furthermore, this test and the studies previously described must be as homogeneous as possible, in order to compare results and provide further evidence for what has been already demonstrated in other research fields. Thus, many key concepts of the theory of affordances need to find an exact counterpart in the framework of this experiment, i.e. a component that reflects it at a behavioural or linguistic level, as will be clarified in what follows. These theoretical assumptions and their experimental transposition will be now examined in turn.

### 2.2.1 The role of visual perception

Most theories on affordances assume that they must be at least visually perceivable (if not already perceived; cf. §1.2.1) and the importance of visual perception has been stressed in many studies, since Gibson's works. In line with theoretical assumptions, almost all experimental researches presented in the previous chapter used visual stimuli, in particular objects' pictures (e.g. Tucker and Ellis 1998; 2001; Symes et al. 2007; Buccino et al. 2009). Therefore, the best way to detect the affordance of graspability is to design a task in which objects are visually presented to participants; in this regard, photographs or images have already been proved to be a well-suited kind of stimulus. However, the purpose of this research is to analyse linguistic behaviour related to affordances, so the test will be built around visual inputs, but will lead to a linguistic output. For these requirements, the experiment is designed as follows: a series of images will be presented to participants and an action description task will be proposed to them.

### 2.2.2 Affordances as an inherently relational concept

Since Gibson, affordances are considered as an inherently relational concept that regards both living beings and the environment in which they act (cf. §1.1.1). This is a basic assumption that should be taken into account in designing the test: this experiment is indeed expected to provide information about how people interact (in the etymological sense of the

word, "to act between") with objects, during the action of grasping. For this reason, while providing descriptions of grasps, participants should have to refer both to how an object can be grasped, according to its physical properties, and to how they would be more likely to grasp it, according to their own abilities.

Since the visual-linguistic test does not require performing any concrete action, it is the instruction given to volunteers that ought to stress this point, so that the close connection between agents and objects could emerge. Even if not collected through practical activities, responses should still reflect the assumption of complementarity stated thus far. For this reason, the adverb *how* will be the keyword of the task: volunteers will be asked to describe how they would grasp a series of objects. Leaving any other specification aside, the word *how* should activate simultaneously both proprioception, i.e. the perception of oneself and the awareness of oneself's abilities, and exteroception, i.e. the perception of objects' salient properties. The linkage between the two types of perception perfectly fits in with the concept of affordances outlined so far. The adverb *how* points two ways, thus focus shifts from the agent-side to the object-side are expected: answers with the focus on the agent-side might be those in which body parts are explicitly mentioned («with my right hand», and the like), whereas when the attention on the object increases or predominates, object's parts might be nominated («by the handle», and the like).

### 2.2.3 Affordances, familiarity with objects and past experience

The way humans interact with objects is deeply influenced by their experience and their familiarity with them. Familiarity also seems to affect the activation of the sensory-motor system during object presentation (§1.3.4). For this reason, most of the cited studies on affordances adopt common everyday objects as stimuli, such as a frying pan, a knife, a mug (e.g. Riddoch et al. 1998; Ellis and Tucker 2000; Tucker and Ellis 1998; 2001; Phillips and Ward 2002; Buccino et al. 2009), while only a few use geometrical shapes (Gentilucci 2002; Symes et al. 2007).

Since using familiar objects in experiments enhances the possibility to observe affordance effects (Pavese and Buxbaum 2002: 562), highly meaningful, concrete and familiar objects will be used as stimuli. A good way to select them is to choose from a list derived from speech corpora (as will be illustrated in more detail in §2.3.2): in this way, only objects that are frequently referred to in speech will be included in this study (cf. §2.3.3).

### 2.2.4 Objects-stimuli

The sample of objects chosen as stimuli must be representative of the entire typology of graspable objects; therefore, variation both in terms of semantic types (artefacts, natural kinds, humans) and in terms of dimensions (big or small) has to be guaranteed, even if the studies presented in Chapter I are only conducted on artefacts and natural kinds.

Also, there is another objectual characteristic that is worth being taken into account: the presence of affording parts. Affording parts play an important role during objects visual processing and recognition: for instance, we have shown that the perception of a cup with a broken handle causes a decrease in motor evoked potentials recorded from hand muscles, compared to those raised by a cup with an intact handle (Buccino et al. 2009). For this reason, artefacts will be divided into two classes: objects with a part that is particularly suited for grasping (e.g. a handled cup), and objects without any part that more than others attracts the grasp (e.g. a tennis ball).

Moreover, since many studies demonstrated that objects with affording parts elicit different motor representations (or different micro-affordances, according to Tucker and Ellis' definition: see §1.3.2) as the orientation of the affording part varies, in the experiment here described, too, some objects will be presented in different orientations (leftward/rightward-oriented, upright/overturned).

## 2.3 Data

The list of objects used as visual stimuli was extracted from the corpus developed within the ImagAct project.[10] It is well worth not only explaining how the objects-stimuli were chosen, but also describing the corpus and how it was created, since it will also be the basis of another experiment that will be presented in Chapetr VI.

[10] Information and references about the project and the ImagAct resource can be retrieved at the URL: http://www.imagact.it.

### 2.3.1 ImagAct: a multimodal and interlinguistic ontology of action

In traditional lexicographic resources, as in most dictionaries, word meanings are generally explained with definitions or with equivalent words (e.g. WordNet). Nowadays, providing visual support for lexical and ontological resources is becoming more and more important. This is demonstrated by a growing number of resources that allow users to access visual content, such as illustrations, pictures, animations or videos (SUMO,[11] Image-net,[12] among others). Parallel to this, the need for retrieving images from large datasets has led to a broad effort to annotate images with semantic content (consider for istance *ImageML*, a model created for the semantic annotation of images representing events, described in Bosque-Gil and Pustejovsky 2015). If on the one hand, images provide a user-friendly representation of concrete entities, on the other hand, videos are effective in presenting verbal meanings, even if this is still an under-investigated issue (videos are rather costly, both in terms of storage space and for their difficult realisation). However, in all existing resources, the hierarchical structure is made of linguistic or conceptual units (e.g. synsets, lexical entries), whereas the multimedial content is linked to them; there is no visual ontology in which semantic and lexical information is dependent on the multimedial content (Frontini et al. 2012). This is exactly the structure of ImagAct.

ImagAct is an ontology exclusively focused on the domain of action verbs (Moneglia et al. 2012; 2013; Panunzi et al. 2014; a detailed and complete description of the resource is provided by Gagliardi 2014). It is organised around short videos, constituting the nodes of the ontology, that represent particular types of actions (e.g. a man taking a glass from a table, or a man taking water from the faucet using a bottle), set in a pragmatically neutral context. Action types are kinds of events conceptually different from each other, that can all be denoted by a single verb (or a class of locally equivalent verbs), but can be also distinguished from one another by considering the ways in which the agent interacts with objects as well as the body movements involved. For instance, we may consider the verb *to take* in the following examples:

[11] http://sigma-01.cim3.net:8080/sigma/Browse.jsp?kb=SUMO

[12] http://www.image-net.org/

1. *John takes a present from a stranger*
2. *John takes Mary the book*
3. *John takes the pot by the handle*
4. *John takes Mary to the station*

These sentences denote completely different types of actions. In (1), the verb *to take* means "to receive, to accept"; but in (2) it means "to bring"; in (3) simply "to grasp" and in (4) "to conduct". Such different action types correspond to events in which agents move and act in different ways, therefore the sets of locally equivalent verbs that in each case may substitute *to take* are different: *John receives a present from a stranger* is acceptable, but *John receives Mary the book* is not; *John brings Mary the book* would rephrase *John takes Mary the book*, but *John brings the pot by the handle* is not equivalent to *John takes the pot by the handle*, etc.

As a matter of fact, humans adopt the same lexical form to denote different types of events. It is evident that action verbs are one of the most difficult parts of the lexicon to learn for L2 learners, because languages segment action concepts in very different ways. For instance, the verb *to fold* may be rendered in different ways in Italian, i.e. *piegare* in *Maria piega la camicia* ("Mary folds her shirt") and *incrociare* in *Maria incrocia le braccia* ("Mary folds her arms"). Similarly, the sentences with the verb *to take* reported above should be translated into Italian using two different verbs, i.e. *prendere* (for the first and the third sentence) and *portare* (for the second and the fourth sentence).

Crucially, the very same verb that can denote a given action type, is also extendible and applicable to all similar events:

1. *Mary folds the shirt/the blanket = Maria piega la coperta/la camicia*
2. *Mary folds her legs/her arms = Maria incrocia le gambe/le braccia*
3. *John takes the children to school/Mary to the station = John porta i bambini a scuola/Maria alla stazione*
4. *John takes the pot/the book/the dish = John prende la tazza/il libro/il piatto*

Each action type in ImagAct is related to a video and each video is in turn linked to a list of Italian, English, Spanish and Chinese verbs that can be used to describe that particular event (other languages are currently being implemented). For each video, there is also a "best example", i.e. a short sentence that exemplifies the use of a verb. Obviously, one verb usually

points to more than one scene, i.e. may be used to denote different types of events (cf. *to fold*).

The greatest innovation, compared to preexisting resources, is that in ImagAct action types, i.e. different kinds of events which a verb may refer to, are not defined by means of glosses or lexical entries, but are represented by videos created as 3D animations. These scenes are associated not only to a verb lemma, but also to an exemplifying sentence. Moreover, ImagAct's videos have been linked to (Ital)WordNet's synsets and a mapping between the two resources has been established and evaluated (Bartolini et al. 2014; De Felice et al. 2014).

The best way to describe how the ontology is constructed and how the web interface works is by looking at a concrete example. The ontology is accessible in different ways to users (for example, L2 learners) with a web interface, through three different sections: Dictionary, Gallery and Compare.

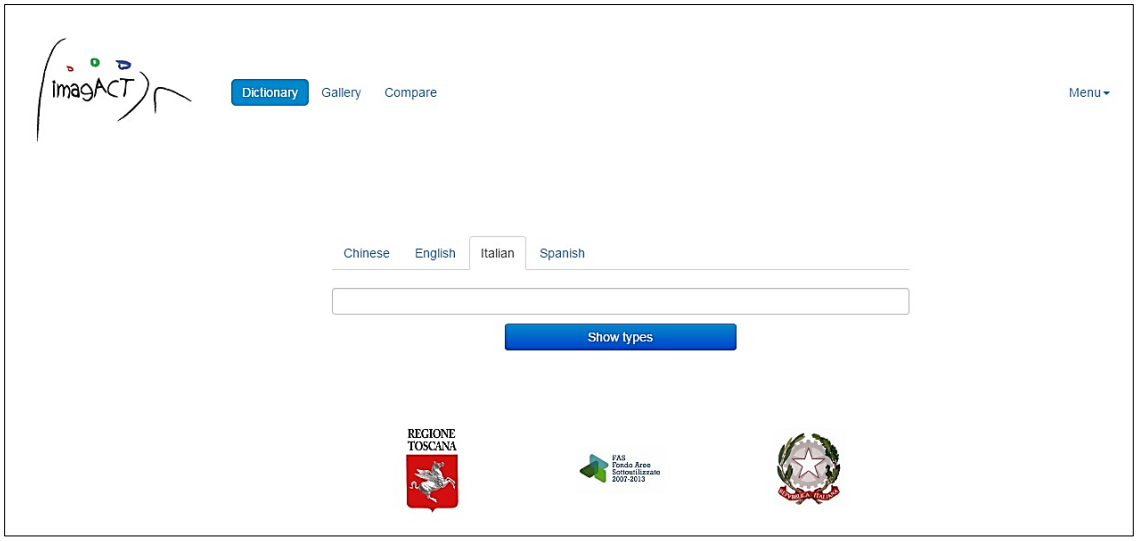



Figure 2.1. Imagact query interface (www.imagact.it).

In the Gallery, the scenes are organised in nine macro-categories (facial expressions, actions referring to the body, movement, modification of the object, deterioration of an object, force on an object, change of location, setting relation among objects, actions in the intersubjective space). For example, among the 313 scenes representing actions that cause the modification of an object, there is the action represented in Fig. 2.2.

Figure 2.2. The action of pushing the ring button (four still images from a single video).

By moving the cursor on the video, the best example appears in the language that has been selected (e.g. *Marta suona il campanello*) and tells us that this particular action is denoted by the Italian verb *suonare*. It is also possible to find possible translations for this verb, in this specific sense.

If users now wanted to know which are the types of action that a verb may describe, they can enter the lemma *suonare* in the Dictionary and find the two action types denoted by the Italian verb. The first one indicates the action of ringing the bell, whereas the second one indicates that of playing an instrument. From this page, again, it is also possible to look at the correct translations (Fig. 2.3).

Scene: 4b8bcda1
Best Example: *Marta suona il campanello*
Italiano: suonare
English: to ring, to sound
Spanish: tocar
Chinese: 按 àn

Scene: bbc50559
Best Example: *Fabio suona il pianoforte*
Italiano: suonare
English: to play
Spanish: tocar, sonar
Chinese: 弹 tán

Figure 2.3. The variation of the Italian verb *suonare* in ImagAct.

This allows us to appreciate the differences among languages: in this case, in both events there is an agent producing a sound, but some languages do not use different verbs to codify these actions (as Italian), whereas some other languages may use distinct lemmas, thus remarking differences among action types.

Finally, if users are interested in comparing the use of different verbs among languages (as well as within a same language), for example between English *to ring* and Italian *suonare*, they can enter the two lemmas in the Compare section: the two verbs cover only one shared action type, only *suonare* can mean "to play an instrument", whereas only *to ring* can mean "to telephone" or "to dial a number" (Fig. 2.4).

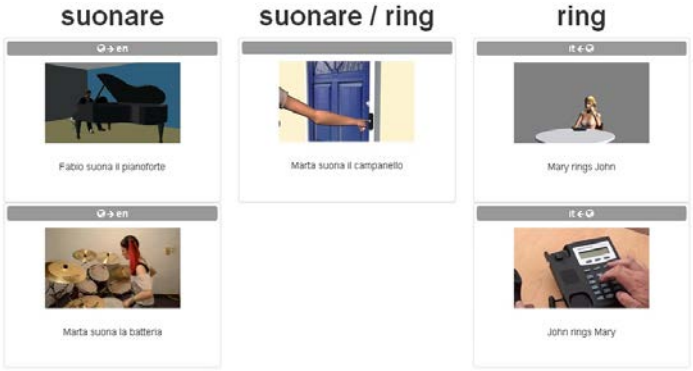



Figure 2.4. An example of the ImagAct Compare function: comparing Italian *suonare* and English *to ring*.

Therefore, the ImagAct ontology can be accessed by lemma, by scene and, furthermore, the usages of two verbs can be compared.

### 2.3.2. The ImagAct corpus

What is of particular interest for the present study is the corpus[13] that is behind this ontology and that is not directly accessible from the web interface yet. Since the first major task of the ImagAct project was to develop an "ontology of action" from Italian and English data (that was later extended to other languages), two corpora specifically focused on high frequency English and Italian action verbs were first created. Action verbs provide the main semantic contribution in sentences and are also the most frequent items in speech (Moneglia and Panunzi, 2007), therefore English and Italian ImagAct corpora have been derived from

[13] For simplicity, we will use the term *corpus* even if we are aware that the "ImagAct corpus", for its peculiarities (as will be shown in what follows), does not fully satisfy the traditional requirements of sampling and representativeness (cf. McEnery and Wilson 2001: 32). On this topic, cf. Gagliardi (2014: 58).

different parts of pre-existing spoken corpora. For English, ImagAct exploited a random sampling of the BNC-Spoken corpus (of around 2 million words). The Italian part of the corpus consists in a collection (of around 1.6 million words) of spontaneous speech corpora available for research, in particular the entire LABLITA Corpus of Adult Spontaneous Spoken Italian (Cresti and Moneglia 2005), the entire Corpus LIP (De Mauro et al. 1993), and part of the Corpus CLIPS[14] (Albano Leoni 2003).

From these collections of texts, the ImagAct corpus was derived in parallel for English and Italian in different steps (Moneglia and Panunzi 2011; Moneglia et al. 2012; Frontini et al. 2012). First of all, all occurrences of high frequency action verbs (more than 1100 Italian or English lemmas) were extracted from spoken corpora with their linguistic context. Then, through a web interface, each context extracted was standardised by expert annotators and reduced to a very simple sentence (present indicative, 3rd person singular, active voice, definite and singular subject, definite object etc.). In doing this, every information not relevant for the action itself was disregarded, in order to create, for each lemma, a list of instances that show which kinds of events people usually refer to, when using these verbs in a real conversation: e.g. a sentence like *then I finally took the red book that was on the table* would have been standardised as *John takes the book*. When there was an implicit (usually anaphoric) reference to an agent or an object, annotators found the correct anchors in the original corpus and made it explicit: for instance, a sentence like *they finally caught them* would have been standardised as *the policeman catches the thieves*.

Annotators then assigned a primary or marked value to all standardised sentences (cf. Fig. 2.5): primary (i.e. proper) sentences occur when an action verb refers to a concrete, physical action, as in *John gave Mary the umbrella*, *John runs to the station*; in marked (i.e. non-proper) sentences, verbs are used in abstract, metaphorical or idiomatic expressions, as in *John gives me a good idea*, *John gives up smoking*, *time is running too fast*, etc. After these preliminary operations, only primary uses of action verbs were considered in building the ImagAct resource.

[14] Only the sub-corpora derived from radio and television broadcasts and part of the dialogues were considered, because they were the most relevant for the purposes of the ImagAct project. For details about the composition of the ImagAct corpus, see Gagliardi (2014: 58 ff.) and Moneglia (2014).

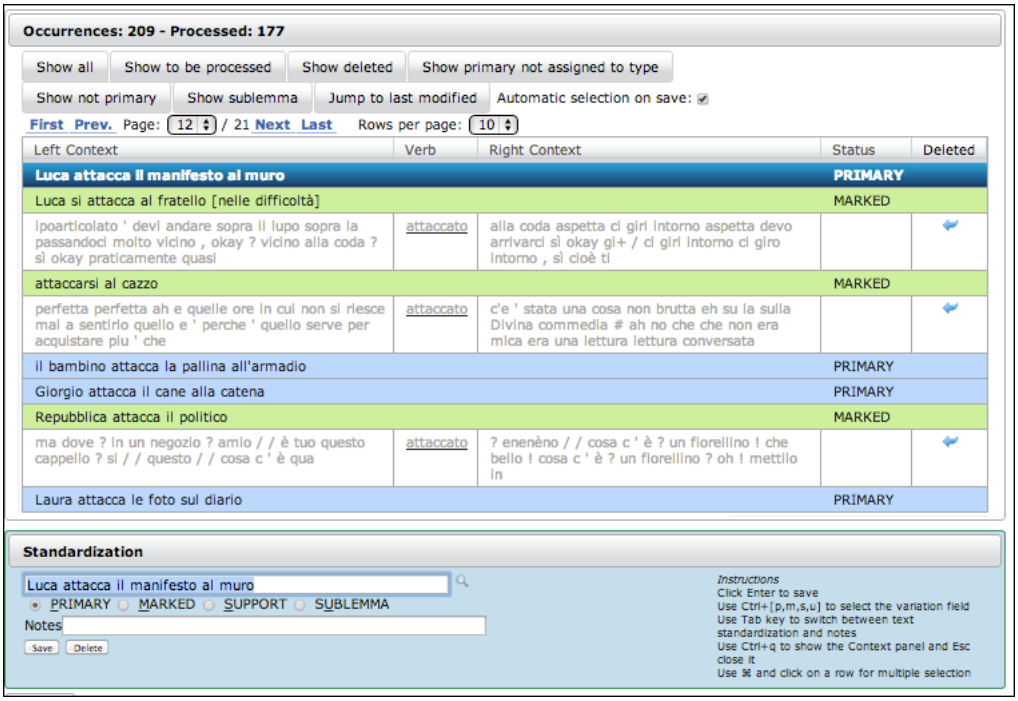



Figure 2.5. Example of the web interface used for the annotation. Assignment of the primary or marked value to the standardised sentences for the verb *attaccare*, "to attach" (from Gagliardi 2014: 76).

The action types introduced in the first part of this paragraph were derived adopting a bottom-up methodology. After standardizing all occurrences of high frequency Italian and English action verbs, sentences were manually clustered on the basis of the similarity among the events they denoted (mainly in terms of body movements and types of participants involved). As it has been already noticed above, it is evident that the type of action described in *Mark takes the book from the shelf* is completely different from that described by *Mark takes Mary home* (see Moneglia et al. 2012; Frontini et al. 2012; Moneglia et al. 2013; Panunzi et al. 2014). Action types, linked to the scenes and to their associated best examples, represent the variation of all action verbs considered and constitute the nodes of the ImagAct ontology of action. Therefore, this ontology not only is inherently inter-linguistic, because it is derived from corpora of different languages through an inductive process, but also takes into account the intra-linguistic and inter-linguistic variation that characterises action verbs in human languages.

### 2.3.3 Graspable objects in ImagAct corpus

The choice of the objects-stimuli used in the experiment was inspired by the information extracted from the ImagAct corpus, which is specifically focused on action verbs as they are used in real conversations and familiar situations. Furthermore, since anaphora is avoided and explicit reference to objects is always restored, this corpus provides direct evidence about the objects typically involved in actional events as they are referred to in speech.

In order to have a list of graspable objects, the verbs pertaining to the semantic field of grasping were selected among the almost 600 Italian verbs present in the ImagAct database: *acchiappare*, *accogliere*, *afferrare*, *cogliere*, *pigliare*, *prendere*, *raccattare*. Then, all their primary instances (tot. 1309) were collected, all direct objects' lemmas were extracted and their frequency in the corpus was calculated.

The objects chosen as stimuli were selected according to different criteria. Since objects-stimuli had to be representative, as far as possible, of the entire typology of graspable objects, they first had to be representative of different semantic categories: artefacts (with or without affording parts), natural kinds, substances and aggregates, and humans. Dimension is another relevant parameter; therefore both large and small objects had to be present (in the specific field of grasping, we consider as "large" a dimension that exceeds the maximum span of a hand). During the selection of the stimuli for the experiment, objects with the highest frequency in the corpus were preferred.

## 2.4 Methods

Before describing in detail the methods adopted for the experiment, two pilot tests that were conducted and proved very useful to define the experimental procedure will be briefly illustrated.

### 2.4.1 Pilot tests

A first test was conducted on four non-expert volunteers. Subjects were presented with the following list of 48 images[15] (the order of appearance is respected):

[15] Visual stimuli will be described in more detail in §2.4.3.2.

| | | | |
|---|---|---|---|
| 1. tennis ball | 13. coffee cup (R, O) | 25. guitar (R) | 37. chair |
| 2. vase | 14. pencil | 26. umbrella | 38. hair dryer (L) |
| 3. plate | 15. stone | 27. microphone (R) | 39. flour |
| 4. pumpkin seeds | 16. scissors (R) | 28. backpack (lateral) | 40. trolley |
| 5. glass | 17. sword (R) | 29. pacifier (L) | 41. rubber boat |
| 6. running man | 18. tea-cup (R, O) | 30. standing girl | 42. microphone (L) |
| 7. ladle (L) | 19. coffee cup (R) | 31. girl bag | 43. coffee cup (L) |
| 8. lighter | 20. apple | 32. water | 44. guitar (L) |
| 9. box | 21. wine bottle | 33. jug (R) | 45. baby |
| 10. tea-cup (R) | 22. sand | 34. tea-cup (L) | 46. pacifier (R) |
| 11. mandarin | 23. hair dryer (R) | 35. backpack | 47. sword (L) |
| 12. banana | 24. jug (L) | 36. ladle (R) | 48. soccer ball |

Table 2.1. Objects-stimuli for pilot tests 1-4 (R: right orientation; L: left orientation; O: overturned).

For each stimulus presented, participants had to answer a short question reported under the figure: «How would you take [X]?», and to another one, presented in the following slide: «How would you take [X] differently?».

Interviews were transcribed in a single file and analysed. The main problem that arose was that three participants out of four gave very short answers, and (at least in one case) with a high degree of underspecification, without considering objects' parts, shape, etc. (a participant answered for almost all objects-stimuli «with my hands»). As a result, answers were also shorter and shorter over time, for a sort of "automatisation" of the answering process. The first object that participants saw might have caused an imprinting effect on following stimuli, so that the simplicity and brevity of the first answer might have influenced the following ones. The first, very generic response («with my hand»), that seems in some sense correct according to the characteristics of the object presented (a tennis ball), could be obviously also applied to objects for which a more precise answer could be given and was maintained unvaried throughout the test.

Therefore, the following test started not with a simple object-stimulus (a small tennis ball), but with a more complex one (a tea-cup), i.e. an object very familiar to subjects and provided with a specific part designed for grasping, in order to elicit a more detailed grasp description. It is more plausible that participants give a more articulated answer at the very beginning of the test (e.g. for the tea-cup), and then give an easier answer after that first response whenever

they could not maintain the same degree of specificity (e.g. for the tennis ball), than the contrary.

Other two important changes were to give the instruction only one time in the first slide of the presentation (not repeating it in every slide), and to express the task not through a direct question («How would you...?»), but indeed through an explicit request of description: participants are asked to see a list of images representing various objects and to describe, in the most detailed way, how would they grasp them (cf. §2.4.4).

Together, such modifications are intended to stimulate longer and more complete answers, because informants do not have to answer a question, but indeed need to organise a discourse in order to formulate a description. While the task gains in naturalness, answers are expected to be more informative.

Finally, the opposition between the backpack presented from a frontal or lateral point of view had not given rise to different answers, so only the backpack from the lateral view (in which more parts were recognisable) was maintained.

With these changes, a second series of interviews was recorded with five subjects. Stimuli were presented as follows:

| | | | |
|---|---|---|---|
| 1. jug (R) | 13. banana | 25. jug (L) | 37. hair dryer (L) |
| 2. tea-cup (R) | 14. coffee cup (R, O) | 26. guitar (R) | 38. flour |
| 3. tennis ball | 15. pencil | 27. umbrella | 39. trolley |
| 4. ladle (L) | 16. stone | 28. microphone (R) | 40. rubber boat |
| 5. vase | 17. scissors (R) | 29. backpack | 41. microphone (L) |
| 6. plate | 18. sword (R) | 30. pacifier (L) | 42. coffee cup (L) |
| 7. pumpkin seeds | 19. tea-cup (R, O) | 31. standing girl | 43. guitar (L) |
| 8. glass | 20. coffee cup (R) | 32. girl bag | 44. baby |
| 9. running man | 21. apple | 33. water | 45. pacifier (R) |
| 10. lighter | 22. wine bottle | 34. tea-cup (L) | 46. sword (L) |
| 11. box | 23. sand | 35. ladle (R) | 47. soccer ball |
| 12. mandarin | 24. hair dryer (R) | 36. chair | |

Table 2.2. Objects-stimuli for pilot tests 5-9 (R: right orientation; L: left orientation; O: overturned).

Volunteers spoke more and more fluidly, because they were free to present their arguments as they preferred, without the constraints imposed by a question-answer structure. They often started the description with a brief statement about the object (e.g.

«oh, well, this, is a large tea-cup, so I think I would take it by its handle...»). When only one answer was given, informants were asked if they could think of taking a given object-stimulus in some other way (or similar requests were formulated). The active participation of the interviewer in the test, far from being a problem, encouraged people to keep on talking about the stimulus. However, with the open question structure, the test was much more time consuming, thus some pictures had to be eliminated. One stimulus (scissors) gave rise to very articulate answers because the related grip was too specific, so it was deleted from the test. Among artefacts with affording parts, the guitar in the two orientations was also eliminated from the stimuli (again, for the high specificity of the grasp it requires), as well as the upright coffee cup (too much similar to the other cups).

After having briefly described these preliminary pilot tests and the necessary refinements, we can now turn to illustrate in detail the methodology adopted in the experiment.

### 2.4.2 The participants and the experimental setting

Thirty participants entered the study. They were all native Italian speakers and undergraduate students of the Faculty of Languages (only one of them was already graduated). They were five males and 25 females. Their ages were between 20 and 27 years, with a mean of 22.6 and a standard deviation[16] (henceforth SD) of 1.52.

All participants were informed about the purpose of the study and gave their consent to the experimental procedure. In order to maintain anonymity, they received an ID number (from one to 30) and filled a form with their name, ID number, age, weight, length, hand dominance and a judgement about their physical strength (from one to five). Finally, the measure of the maximum span of their dominant hand was also recorded in centimetres. The participation in the experiment was voluntary and free, but each informant received a gadget.

The interviews were audio/video recorded at the Phonetics Laboratory of the University of Pisa (Department of Linguistics; Pisa, Italy). Each session involved only two people at a

[16] Standard deviation (σ, or SD) is a measure of variability used in statistics to express how much the data values are dispersed from the mean. A high standard deviation indicates that the data are widely spread over a wide range of values, whereas a low standard deviation shows that the data are concentrated closely around the mean of the data values.

time, the interviewer (Irene De Felice) and one student. During the experimental session, volunteers were seated on a comfortable chair in front of a PC monitor, placed at a distance of about 60 cm from their eyes, whereas the interviewer was seated on their left. A video camera was placed on the left of the PC monitor, oriented towards the participants so that both the monitor and the informant were framed (Figg. 2.6, 2.7).

Figure 2.6. The experimental setting.

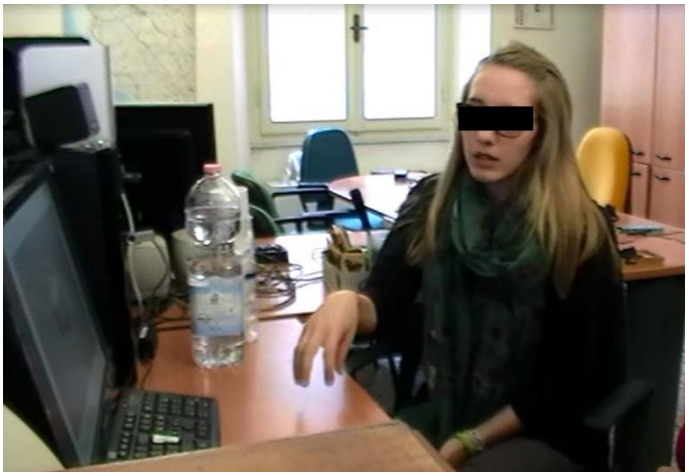


Figure 2.7. The recorded interviews: an example.

### 2.4.3 The visual stimuli

#### 2.4.3.1 Objects-stimuli

Objects used as stimuli for the experiment were all familiar objects, for the reasons exposed in the previous paragraphs. Moreover, they cover the whole range of graspable objects, according to different parameters. In particular, humans, substances and aggregates are also used as stimuli, whereas none of the researches cited in the previous chapter takes into account stimuli different from geometrical shapes, natural kinds and artefacts.

As already mentioned, the ImagAct corpus (cf. §2.3.2) proved very useful for the creation of the list of objects. From its Italian part, all lemmas that occurred as a direct object of verbs denoting actions of grasp were extracted and sorted by frequency. Among these lemmas, stimuli were chosen mainly according to two parameters, dimension (large/small) and semantic type: artefacts; natural kinds; substances and aggregates; humans. Among artefacts, eight stimuli had no parts specifically designed for grasping, whereas the remaining 14 have at least one well-identifiable and prominent part that could attract the grasp.

| |
|---|
| *Artefacts - with affording parts* (23 stimuli): chair; coffee cup (O); girl bag; hairdryer (L/R); jug (L/R); ladle (L/R); microphone (L/R); pacifier (L/R); rubber boat; backpack; sword (L/R); tea-cup (L/R/OL/OR); trolley; umbrella. |
| *Artefacts - without affording parts* (eight stimuli): box; glass; lighter; pencil; plate; soccer ball; tennis ball; vase. |
| *Natural kinds* (four stimuli): apple; banana; mandarin; stone. |
| *Substances and aggregates* (four stimuli): flour; pumpkin seeds; sand; water. |
| *Humans* (three stimuli): baby; man; girl. |

Table 2.3. Objects-stimuli adopted in the experiment (R: right orientation; L: left orientation; O: overturned).

All these stimuli were in the list of objects extracted from ImagAct, except the mandarin, which was added as an example of a small fruit, and the pumpkin seeds, which occupy an intermediate position between two different categories (see *infra*).

### 2.4.3.2 Stimuli presentation

The visual stimuli adopted in the experiment consisted of 42 pictures of concrete entities. They all were photographs of real objects taken against a white background: 22 of them were taken by the candidate (e.g. Fig. 2.8), while the others are public domain images free of copyrights retrieved from the web (e.g. Fig. 2.9).

Figure 2.8. Examples of objects-stimuli: photographs taken by the candidate (rightward-oriented hairdryer, banana, plate, rightward-oriented ladle).

Figure 2.9. Examples of objects-stimuli: public domain images (rightward-oriented pacifier, rightward-oriented sword, chair, and child).

Only two pictures (the running man and the standing girl) maintained a realistic background, i.e. a city setting, in order to make the requested task easier for the participants, who in this way could more easily imagine a real context in which they might have to grasp a person. A complete list of the images used as stimuli can be found in Appendix A.

In the category of artefacts with affording parts, there are some distinctions to be made between different objects. Whereas most of them have a real handle, i.e. only one protruding

part that is specifically designed for facilitating the grasp (as for the umbrella, the cup, the hairdryer), there are also objects that have more than one part that could possibly be grasped. For instance, the rubber boat has two handles, but it also has a rope surrounding it, as well as some plastic rings where the rope is inserted.

The chair as included in the list of artefacts with affording parts even if it does not have a specific part designed for grasping, because, with comparison to artefacts without affording parts, it is characterised by the presence of more than one discrete and visually distinguishable part, which could attract the grasp. Moreover, these parts are not equivalent to one another: in particular, the back of the kind of chair chosen as visual stimulus is the part probably most suited for grasping, because it reaches the height of human hands. The seat and especially the legs are more difficult to grasp.

In the category of substances and aggregates, four different kinds of stimuli are grouped. The water represents a liquid, a canonical example of a substance; water's minimal elements are continuous and visually undistinguishable. The sand and the flour are aggregates of minimal particles that can be visually distinguished (especially the grains of the sand), but still are very small and humans do not usually interact with them. On the contrary, pumpkin seeds are presented as an aggregate, because in the picture they are collected in a mound (cf. Appendix A), however their minimal elements are larger with respect, for example, to a grain of sand, therefore are perceptively more relevant and clearly distinguishable. The single seeds are also more accessible and humans may interact with them (for example, when eating the seeds one by one).

From a morphosyntactical point of view, the words that denote these four entities in Italian are diverse: the words for water, flour and sand are mass nouns (for example, when combined with the numeral *due*, "two", as in *due acque*, *due farine*, *due sabbie*, they may only refer to two different kinds of water, flour and sand).[17] On the contrary, the word *seme*, "seed" is a count noun (e.g. *due semi*, "two seeds"), but it is worth noting that in Italian both count and mass terms may refer to aggregates, intended as collections of relatively small and homogeneous entities (Middleton et al. 2004).[18] Therefore, the mound of pumpkin seeds can be considered as intermediate between two different categories, distinguished both

[17] Cf. Chierchia (1998).

[18] For example *riso*, "rice" is a mass noun. There are languages in which also small fruits and vegetables are mass nouns. For instance, Wierzbicka (1988: 313) reports that the Russian words for peas and beans (*gorox, gorošek, fasol'*) are mass nouns, just like the words for rice and flour (*ris, muka*) are.

perceptively and linguistically: that of individual objects (such as the pencil, denoted in Italian by count nouns) and that of granular aggregates (such as the sand, mostly denoted in Italian by mass nouns; cf. Clausen et al. 2010).[19] Since the seeds are very small and usually occur together, they are presented them in a mound in order to investigate which type of grasp they afford (i.e. more similar to that of the sand or to that of the lighter).

### 2.4.4 The task

The PowerPoint presentation displayed on the computer monitor started with a white slide with black writing that explained the task. Participants were asked to see a list of images representing various objects and to describe, in the most detailed way, how would they have grasped them («Osserva bene gli oggetti che vedrai rappresentati; quindi, descrivi nella maniera più dettagliata possibile come prenderesti gli oggetti»). Moreover, it was explicitly specified by the interviewer that the description had to be performed verbally, and that gestures and pointing to the pc monitor should not have substituted it. When only one answer was provided, informants were asked if they could think of grasping a given object in some other way (otherwise, similar requests were formulated).

Objects' figures were then presented in the following order, alternating with empty white slides (see Appendix A for figures):

[19] Pumpkin seeds might be considered, following Wierzbicka (1988: 338), "pluralia mostly", because they are possible to count but usually are not counted (cf. Eng. *peas*).

| | | |
|---|---|---|
| 1. jug (right) | 15. coffee cup (overturned) | 29. water |
| 2. tea-cup (right) | 16. stone | 30. tea-cup (left) |
| 3. tennis ball | 17. sword (right) | 31. ladle (right) |
| 4. ladle (left) | 18. tea-cup (overturned, right) | 32. chair |
| 5. vase | 19. apple | 33. hairdryer (left) |
| 6. plate | 20. sand | 34. flour |
| 7. pumpkin seeds | 21. hairdryer (right) | 35. trolley |
| 8. glass | 22. jug (left) | 36. rubber boat |
| 9. running man | 23. umbrella | 37. microphone (left) |
| 10. lighter | 24. microphone (right) | 38. baby |
| 11. box | 25. backpack | 39. pacifier (right) |
| 12. mandarin | 26. pacifier (left) | 40. sword (left) |
| 13. banana | 27. standing girl | 41. tea-cup (overturned, left) |
| 14. pencil | 28. girl bag | 42. soccer ball |

Table 2.4. Objects-stimuli used in the experiment (right: rightward orientation; left: leftward orientation).

### 2.4.5 Orthographic transcription

The duration of the recordings ranges from a minimum of 00:08:07 to a maximum of 00:21:08 per participant, with a mean of 00:13:43 minutes (SD=00:03:54). The speech corpus amounts to more than six hours of speech (06:51:25). Data from audio and video recordings are transcribed in a CHAT format (MacWhinney 2000). The CHAT transcription system is nowadays one of the best-known methods to generate transcripts from oral materials, mainly because it allows researchers to notate only what they consider relevant to their purposes. For instance, it is possible to adopt an advanced and very fine-grained transcription, that also considers proxemic activities, gestures, facial expressions, paralinguistic materials, hesitations, phonological and prosodic features, and many other aspects of a human verbal interaction; but it is also possible to choose a very basic format for transcription and coding that results in a broad transcription (minCHAT). The conventions and principles of the CHAT transcription system are accurately described in the CHAT manual.[20]

[20] http://childes.talkbank.org/manuals/CHAT.pdf

All CHAT files are made of three major components, i.e. the file headers, the main tier and the dependent tiers; these three parts are designed to contain different material and for this reason each one of them has its own notations. There is a rich set of transcription markers, as the manual illustrates, but the symbols and codes adopted in this study are those presented in Moneglia (2005) adopted for the creation of the C-ORAL-ROM corpus. The next paragraphs provide a detailed description of which features were considered relevant and how they were notated in the transcripts.

#### 2.4.5.1 Metadata

Metadata contain all possible information about the recorded session and are included in transcripts as a set of headers lines. For instance, these file headers are the parts of a CHAT file that provide information about the setting and the participants of the communicative event that is being transcribed. Headers are lines of text generally inserted in a fixed order that are introduced by the "@" sign followed by the header name. Few headers are composed only by the header name: these are called *bare headers*, notably the @Begin/@End headers, that are obligatorily used to mark the starting/ending point of the transcript; but most headers also require entries, that specify the value of the header name. Necessary headers of this type are the @Languages and @Participants headers, which provide information about the main language of the dialogue and about the participants (identified by a three letters ID, eventually a name, and a role). Another compulsory header is the @ID header; this line may encode a richer set of data about the participants. It is worth noting that researchers can adopt many other possible headers, according to the specificities of their study.

All CHAT transcripts of the experiment share a common headers structure. Obligatory headers are the following:

@Begin
@Languages: ita
@Participants: 001 Informant, IDF Interviewer
@ID: ita|001|Informant|23|male|right-handed|
@ID: ita| IDF|Interviewer|27|female|right-handed|
…
@End

Each file contains the transcript of a single interview, therefore participants are always two: one is the *informant*, identified with a progressive number (from 001 to 030); the other one is the *interviewer*, identified by the initials IDF (Irene De Felice).

In the @ID line, other information about the informant are manually inserted: age, gender, hand dominance. In the initial part of the file, also, the date and location of the interview are specified and a brief description of the experimental setting is given:

@Location: Pisa, Italy
@Date: 24-MAR-2014
@Room Layout: Phonetics Lab, University of Pisa; the informant is sitting in front of a computer monitor, the interviewer is sitting on his left.; the interview is recorded by a video camera placed on the left, oriented towards the informant.

All transcripts are linked to a video in a *.mov file format: its name coincides with the name of the informant of the dialogue recorded.

@Media: 001, video
@Time Duration: 00:09:13

Two header lines give information about the transcription itself: all data are transcribed by a single person; only phenomena relevant for the present study are notated, so the transcription is coarse and users are warned about some of its peculiarities. The @Situation header precedes the transcription and describes the task being performed:

@Transcription: coarse
@Transcriber: Irene De Felice
@Warning: overlapping was not accurately transcribed.
@Situation: the informant is presented with a list of pictures of objects on the computer monitor and s/he is asked to describe how would s/he grasp them.

After the @Situation heading, a list of @G headings signals the beginning of *gems* and further subdivides the file. Each gem is named with the names of the objects involved in the task and identifies the part in the transcript that corresponds to the part of the dialogue that

follows the presentation of a specific stimulus. @G headers also improve the readability of the transcript: each @G line roughly corresponds to a change of the image presented to the informant on the computer monitor.

@G: jug (right orientation)
[transcription]
@G: tea-cup (right orientation)
[transcription]
@G: soccer ball
[transcription]

#### 2.4.5.2 Main tier

The main tier contains the transcription of what the participants actually said.

Different from the written language, the textual organisation of spoken language is built on utterances, which are the linguistic units of analysis of speech. Defining what is an utterance is not a trivial issue (cf. Moneglia 2005: 15-16) and it is not clear how to distinguish boundaries among different utterances. From a perceptual point of view, they are not only natural units of speech bounded by the speaker's silence, even if in many cases they begin and end with a pause.

Cresti (2000) assumes that a systematic correspondence exists between the information pattern of the utterance and its prosodic patterns. Thanks to this correspondence, we are able to segment the speech flow into utterances by perceiving different prosodic patterns within it (Cresti 1994; 2000).

In this view, utterances are defined as the minimal linguistic units such that they allow a pragmatic interpretation in the world (Cresti 2000): they are autonomous and concluded from a pragmatic point of view (Quirk et al. 1985) and are the linguistic counterparts of communicative acts (cf. Austin 1962). Intonation plays a fundamental role in identifying a speech act: all utterances have a profile of terminal intonation (Karcevsky 1931; Crystal 1975).

Since a single utterance can also be further articulated into various tone units (*simple* vs. *compound utterances*), there are different kinds of prosodic boundaries: on the one hand, those that terminate a sequence of tone units and mark the accomplishment of an illocutionary act (i.e. *terminal breaks*); on the other hand, those that only signal the flow of a same prosodic

programme within a compound utterance (i.e. *non-terminal breaks*).[21] It is worth noting that the annotation of prosodic breaks does not constitute a "transcription of prosody": it does not deal with modelling F0 movements, nor with marking lengthening or intensity peaks.

In line with these theoretical assumptions, the transcription of the interviews recorded in this study is performed in the CHAT standard format and follows the specifications described in Moneglia (2005), adopted for the creation of the C-ORAL-ROM corpus.

The main tier is the core of the transcript and contains the speech transcription. Each line of this part begins with an asterisk, followed by a three letters speaker ID, a colon and a tab, and finally the orthographic transcription of the speech (in Roman alphabet, with words separated by a single space).

The segmentation of the speech flow into utterances is performed through perceptive judgements and distinguishes two types of prosodic breaks. It is useful to quote the passage in which different kinds of prosodic breaks are defined (cited from Moneglia 2005: 17):

> *Prosodic break*: perceptively relevant prosodic variation in the speech continuum such as to cause the parsing of the continuum into discrete prosodic units.
> *Terminal prosodic break*: given a sequence of one or more prosodic units, a prosodic break is considered terminal if a competent speaker assigns to it, according to his perception, the quality of concluding the sequence.
> *Non-terminal prosodic break*: given a sequence of one or more prosodic units, a prosodic break is considered non-terminal if a competent speaker assigns to it, according to his perception, the quality of being non-conclusive.

All utterances are assumed to end with a perceptively relevant prosodic break that has a terminal value (i.e. terminal prosodic break) and is marked with the double slash (//). Non-terminal prosodic parsing is signalled with a simple slash (/).

| | |
|---|---|
| @G: | coffee cup (u/d) |
| *007: | questa la prenderei / sempre / mettendo la mano / come se dovessi afferrare una pallina // |
| *IDF: | mhmh// |

[21] A similar distinction between strong and weak breaks is found in Buhmann et al. (2002).

*007: oppure cercherei / di mettere / il dito nell'occhiello // per poterla poi rigirare //
*IDF: ok//

There are also three more types of terminal breaks that respectively mark interrogative utterances (ending with a "?" tag), intentionally suspended utterances ("…") and interrupted utterances ("+"). All these symbols are described in Moneglia (2005: 26 ff.).

*002: perché se magari devo fermarlo per chiamarlo / non lo so / per un +
*IDF: mi basta che lo prenda in qualche modo //
*002: per un braccio // o per la maglia // o per [/] forse per il polso //

The transcription method here described (cf. Moneglia 2005 for further details) is an implementation of the CHAT format, but many relevant differences between the two systems can be observed. For instance, in the traditional CHAT format (MacWhinney 2000), the comma and the semicolon may mark syntactic pauses and light intonational drops, as in the common practice. Moreover, the period and the exclamation point, together with the question mark, are the basic utterance terminators in CHAT: every sentence must end with one of these punctuation signs, which are used as in the conventional punctuation system. These symbols are not adopted in this transcription.

The [/] symbol is used to mark a retracting break, a special type of non-terminal break. Retractions are the most frequent fragmentation phenomenon in spontaneous speech and they are generally associated to speaker's hesitations. Contrary to interruptions, retracting is usually accompanied by the complete or partial repetition of the preceding linguistic material and by a prosodic break (Moneglia 2005: 27). In this regard, both traditional CHAT symbols [/] and [//], that respectively mark complete or partial repetition of a preceding linguistic material, are expressed just by one symbol [/]. Errors followed by a correction are not distinguished from other retractions. The symbol [/] is inserted in the position where the restart begins.

*002: mah io lo prenderei per / non lo so / eh / per [/] per chiedergli qualcosa//
*002: questo / ci berrei [/] ci berrei//
*IDF: anche per altre &fa [/] per fare altre cose?

As exemplified in the last line, all fragmentary words are immediately preceded by the ampersand symbol "&".

In the transcripts collected for this study, the sequence "yyyy" indicates linguistic material that is not relevant for the analysis, therefore is not transcribed; "xxx" indicates a sequence of unintelligible words; "hhh" is used to signal laughing. The angle brackets "<" and ">" symbols can be used to identify which part of the utterance (if it is not restricted to the word that immediately precedes the string) is affected by the paralinguistic event. Other types of behaviour (such as crying, yelling, coughing etc.) are not found or are deliberately not notated in the transcripts. The following table summarises the symbols adopted in the transcription.

| Symbol | Description |
|---|---|
| / | Non conclusive prosodic break |
| // | Conclusive prosodic break |
| + | Conclusive prosodic break such that the utterance is interrupted by the listener or by the speaker himself |
| … | Conclusive prosodic break such that the utterance is left intentionally suspended by the speaker |
| ? | Conclusive prosodic break such that the utterance has an interrogative value |
| [/] | Non-conclusive prosodic break caused by a false start (retracting phenomena with complete or partial repetitions) |
| & | Speech fragments |
| hhh | Paralinguistic elements (laughing) |
| xxx | Incomprehensible word or sequence of words |
| yyyy | Non transcribed audio signal |

Table 2.5. Symbols adopted in transcripts (adapted from Moneglia 2005).

### 2.4.5.3 Dependent tiers

Dependent tiers are inserted below a main line and contain various material of interest to the researcher (such as comments, event descriptions, etc.). Here, only four dependent tiers are adopted (for a complete list of the possibilities offered by CHAT, see MacWhinney 2000; cf. also Moneglia 2005: 37 ff.):

%com:

The comment tier is a general-purpose line, in which the transcriber can insert notations of various kind relevant for the study or for the comprehension of the transcript.

%tim:

In the present study, this tier is used to notate the duration of the untranscribed parts of the transcript (see *infra*).

%exp:

This level, too, is used in connection with the "yyyy" string in the main line and contains information about what happened during the untranscribed parts of the interview, e.g.:

| | |
|---|---|
| *IDF: | yyyy// |
| %tim: | 15:20:22-17:34:21 |
| %exp: | the interviewer talks with the informant about arguments which are not relevant for the present study. |

To this day, the first phase of the orthographic transcription of the interviews, i.e. the transcription of the speech production of informants, is completed; however, the notation of prosodic breaks according to the specifications described in §2.4.5.2 is still ongoing. As soon as the work will be completed, the corpus will be freely available for research.

After having described the methodology adopted in the experiment and the transcription system adopted, we can now turn to the analysis of the data recorded from the interviews.

# Chapter III

# From Hands to Handles: How Objects Affect Linguistic Production

This chapter describes a linguistic analysis conducted on the transcripts of the interviews recorded during the experiments. First, the purpose of this analysis and the annotation of the transcripts will be illustrated (§3.1). Then, the results of the annotation will be presented in the following paragraphs, which will be dedicated to different semantic classes of visual stimuli, i.e. artefacts (§3.2), further subdivided into artefacts without affording parts, artefacts with affording parts, and artefacts with affording parts that were presented with different orientation; humans (§3.3); natural kinds (§3.4); and substances and aggregates (§3.5). Finally, general results will be presented (§3.6) and briefly discussed (§3.7).

## 3.1 The analysis of the linguistic descriptions of grasps

The assumption that guides this first analysis of the transcripts is that the linguistic descriptions of grasps provided by informants, even with a superficial analysis, may tell us something about the affordance of grasping.

As already mentioned, the concept of affordances is an inherently relational one, which takes into account both the agents' abilities and the objects' physical properties («An affordance points both ways, to the environment and to the observer»: Gibson 1979: 129).

Is it possible to observe in grasp descriptions shifts of attention from the object side to the subject side, or vice versa, with relation to the presentation of different stimuli? Are semantic categories of stimuli related to differences in lexical choices? Moreover, neurophysiological and behavioural studies demonstrate that some characteristics of the object are more able than others to activate the sensory-motor system. Are objects' salient parts (affording properties) more likely to be named than non-salient parts? These are the main research questions that will be addressed in what follows.

### 3.1.1 Purposes and methods

This chapter describes an analysis of the transcripts focussed on the explicit mention of the effector or the target of the grasp, which are defined as follows:

- the *effector* of the grasp is the entity that is linguistically presented as the one that comes in contact with the object;
- the *target* of the grasp is the part of the object-stimulus where the contact with the effector is described to occur.

The purpose of this study is to investigate if, and in case how, the number of references to the effector and the target of the grasp varies depending on different types of stimuli; in other words, if the presence of the effector or the target of the grasp in the event linguistically described is modulated by the presentation of different kinds of graspable objects.[22]

[22] Therefore, in this analysis, any information related to the modality of the grasping and in particular to the

Since we are concerned with affordances, which are primarily related to the most immediate and direct interaction with the object (affordances are automatically raised by object visual perception), only the participant's first grasp description for each stimulus was considered. Therefore, all first descriptions provided by the 30 informants for the 42 objects-stimuli were collected (tot. 1260, approximately 19000 words). The segmentation of speech into different parts corresponding to single descriptions was a simple task, because in most cases they are clearly delimitated by syntactic and/or utterance boundaries. Such grasp descriptions were collected and imported in a single file. The methodology adopted to conduct the analysis of this material can be articulated in two parts, which will be described in the two next paragraphs.

### 3.1.2 Extraction of target-related and effector-related words

First of all, all lexical words used to refer to the effector or to the target of the grasp (defined as in the previous paragraph) were manually extracted from each grasp description. For instance, consider the following three descriptions, produced for the microphone, the overturned tea-cup, and the box respectively:

*027: allora // questo / lo prenderei con la mano sinistra // e basta //
*027: la prenderei dalla parte superiore / con la destra //
*010: con le mani / ai lati //

In the first description, only the effector of the grasp is indicated (*con la mano sinistra*) and two words, *mano*, "hand" and *sinistra*, "left", are extracted. In the second one, we have both a reference to the effector, i.e. *con la destra* (effector-related words: *destra*, "right hand") and a reference to the target of the grasp, i.e. *dalla parte superiore* (target-related words: *parte*, "part", *superiore*, "upper"). In the third one, the informant referred to the effector of the grasp (*mani*, "hands") as well as to the target (*lati*, "sides").

Effector-related words denote the entity that is linguistically presented as the effector of the grasp, therefore in most cases are referred to the hand, such as *mano* ("hand"), *destra*

verbal lexicon (used to describe the kind of action performed by the effector on the target) is not taken into account.

("right hand"), *sinistra* ("left hand"), or hand's parts (such as *dito*, "finger", *pollice*, "thumb", *indice*, "index", etc.). However, sometimes participants referred to different body parts (e.g. the arm, the mouth) or, very rarely, even to an instrument (e.g. a glass, used for water).

On the other hand, we have defined the target of the grasp as the part of the object-stimulus towards which the action is directed and where the contact with the effector is described to occur. Therefore, target-related expressions generally denote a specific part of the stimulus presented (e.g. *manico*, *impugnatura*, "handle");[23] however, informants may also use nouns, adjectives and adverbs pertaining to the visuo-spatial domain, used to indicate the specific part of an object with which the effector comes in contact (e.g. *parte superiore*, "upper part"; *lato destro*, "right side"; *centro*, "centre"; *base*, "base").

A detailed analysis of the target-related and effector-related words extracted from transcripts will be presented in Chapter IV (in particular, cf. §4.1, §4.2), focussed on how the reference to the effector or the target of the grasp is actually made by informants. As already stated, the purpose of this first analysis is only to investigate if the presence of an explicit reference to the *effector* of the grasp (the entity that is linguistically presented as the one that comes in contact with the object) or to the *target* of the grasp (the part of an object-stimulus where the contact with the effector is described to occur) varies with relation to different types of stimuli.

### 3.1.3 The classification of grasp descriptions

Each grasp description was automatically classified according to the presence of one or more words denoting the target or the effector of the grasp. If more than one target or effector-related word is found in a single description, they are counted as one. This is because even if the informants repeat information or reformulate their discourse, the reference to the effector (or to the target) of the grasp is always one and unvaried, as the following example shows:

[23] Participants sometimes used diminutives or odd expressions to identify objects' parts. For instance, a student native to Sardinia, for almost all objects-stimuli with affording parts, named the word *manica* (Engl. "sleeve"), instead of *manico* (Engl. "handle"). This misleading confusion of lexical gender is probably due to an influence of the Sardinian language, in which *mániga* or *mánica* (feminine) is the equivalent of *manico* (Puddu 2000).

*011: mh / dal [/] dal manico dell'oggetto // dall'impugnatura / con una mano //

In this excerpt, one word related to the effector of the grasp (*mano*) and two different words denoting the target of the grasp (*manico*, *impugnatura*) are extracted from the transcript. However, the description actually shows just one reference to the effector and one to the target: *manico* ("handle") and *impugnatura* ("grip") denote exactly the same object's part. This example is just to demonstrate that it is not possible to find, *in a single description*, the reference to two different effectors of the grasp, as well as to two different targets. It is only possible to find a more precise definition of the effector or the target already mentioned (e.g. con la mano // la mano destra),[24] or simply a reformulation (as in the example cited above, for *manico* and *impugnatura*).[25] At this level of analysis, the maximum of information that a single, complete description can provide is just *one* reference to the effector and *one* reference to the target, even if both the effector and the target of the grasp are referred to by more than one word. As a consequence, for each object-stimulus we may have a maximum of 30 references to the effector and 30 references to the target (tot. 60 references), in the event that 100% of the participants named both the effector and the target of the grasp in their first description provided for the stimulus.

Table 3.1 presents an example of the classification, taken from the transcripts for the leftward-oriented jug:

[24] "with my hand // my right hand".

[25] Descripions involving a dinsjunctive conjunction (such as: questa dal manico // con la destra / o con la sinistra) were split into two different descriptions (in this example, only *con la destra* is considered to refer to the effector of the frst grasp description).

| Transcript | Effector | Target |
|---|---|---|
| *018: allora // questa / con la *mano sinistra* // potrei afferrare il [/] il *manico* // | 1 | 1 |
| *015: eh / questo // dal *manico* // | 0 | 1 |
| *020: questa / sempre con [/] dal *manico* // però con la *mano sinistra* / mi verrebbe da prenderla // | 1 | 1 |
| *013: con le *mani* // | 1 | 0 |

Table 3.1. Example of the classification of the grasp descriptions: presence of a reference to either the effector, or to the target of the grasp, or to both. Words referred to the effector or to the target of the grasp are in italic font.

The results of this classification are analysed by considering different semantic categories in turn.

## 3.2 Artefacts

Within this category, linguistic data from artefacts without affording parts will be first analysed (§3.2.1). We will then turn to artefacts with affording parts, presented with or without different orientation (§3.2.2, §3.2.3). Finally, the two groups of stimuli will be compared and some further considerations will be made (§3.2.4). This differentiation is necessary not only because this is the most numerous class of stimuli, but also because some intra-category differences emerge, as the following paragraphs will illustrate.

### 3.2.1 Artefacts without affording parts

This first class of artefacts collects the eight visual stimuli representing objects without affording parts. Figure 3.1 shows the number of references to the effector and to the target of the grasp ($x$ axis) contained in the descriptions provided for different objects-stimuli ($y$ axis). The total number of references collected from the 30 transcripts are shown beside the bar labels.

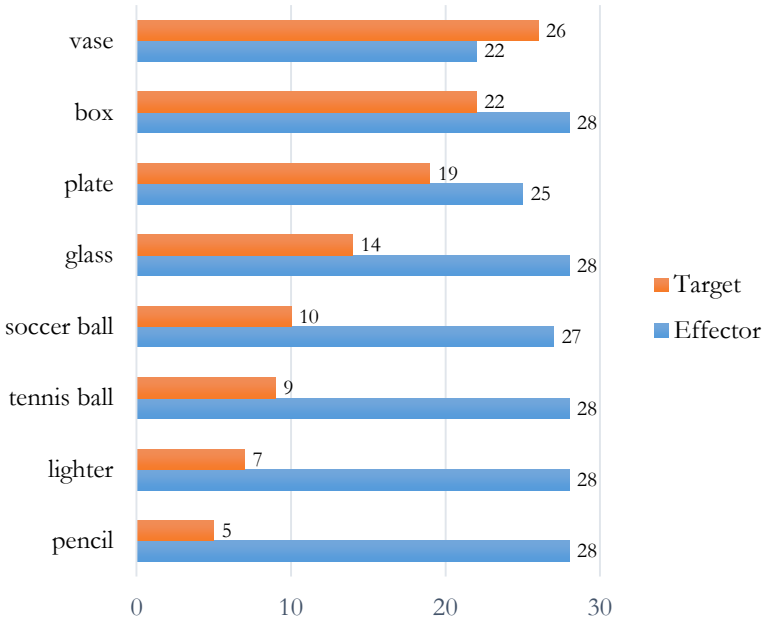



Figure 3.1. References to the effector (blue bars, tot. 214) and references to the target (red bars, tot. 112) found in grasp descriptions for artefacts without affording parts.

The distribution shown by the graph is clearly not due to chance; stimuli are different from one another, with respect to the presence of a reference to either the effector or the target of the grasp ($\chi^2$ (7, $N$=326)=23.082, $p$<0.01).[26]

The number of descriptions that contain an explicit mention of the effector is high for most visual stimuli (mean=26.75; SD=2.19) and does not show great variation, ranging from a minimum of 22 to a maximum of 28 descriptions per stimulus. This indicates that each one of the thirty informants, for most objects-stimuli pertaining to this category, mentioned the effector of the grasp. On the other hand, references to the target of the grasp (mean=14; SD=7.6) vary a lot: only five out of 30 descriptions collected for the pencil contain a mention to the target of the grasp, whereas 26 descriptions for the vase contain this kind of information.

Although the visual stimuli in this category have all been classified as artefacts without affording parts, it seems that at least for some of them the lexical choices made by informants

[26] Statistical tests were performed with the SPSS (Statistical Package for Social Science) software for Windows (SPSS., Inc., Chicago, IL, USA), for which cf. Field (2013). Chi-square statistics are reported with degrees of freedom and sample size in parentheses, the Pearson chi-square value, and the significance level.

reveal the high salience of some of the object's parts, as if they afforded action more than others. In particular, the vase has far more mentions of the object than expected (standardised residual[27]= 2.3), as Table 3.2 shows.

| | | vase | box | plate | glass | soccer ball | tennis ball | lighter | pencil | tot. |
|---|---|---|---|---|---|---|---|---|---|---|
| **Effector** | frequency | 22 | 28 | 25 | 28 | 27 | 28 | 28 | 28 | 214 |
| | std. residual | -1.7 | -0.8 | -0.7 | 0.1 | 0.6 | 0.8 | 1 | 1.4 | |
| **Target** | frequency | 26 | 22 | 19 | 14 | 10 | 9 | 7 | 5 | 112 |
| | std. residual | 2.3 | 1.2 | 1 | -0.1 | -0.8 | -1 | -1.4 | -1.9 | |

Table 3.2. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for artefacts without affording parts.

This is not surprising. Among the stimuli in this category, the vase is the object that, more than others, is constituted by identifiable parts, some of which are particularly suitable for grasping, such as the neck or the edge. These are frequently named by informants. The most frequent target-related word found in transcripts is *collo* ("neck", eight occurrences), named as a possible target for the grasp. It is worth noting that the vase is also the only stimulus for which the frequency of references to the target (26) exceeds the frequency of references to the effector (22).

### 3.2.2 Artefacts with affording parts (without different orientation)

For artefacts with affording parts, first we report data from visual stimuli not presented with different horizontal orientation. For this category, the difference among different stimuli

[27] When using the chi-square of contingency tables with more than two rows or columns and a significant value of $\chi^2$ is obtained, it is possible to consider standardised residuals to know which cell or cells are most responsible for the significance of the test and this knowledge helps in interpreting the results. In particular, when the standardised residual of a cell exceeds the value of ±1.96 (corresponding to an alpha of 0.05) or ±2.58 (corresponding to an alpha of 0.01), the cell deviates from its theoretical value enough to be regarded as an "abnormal" cell, which contributed to the significance of the chi-square test. Standardised residuals with a positive or negative value indicate that the cell is, respectively, over-represented or under-represented in the actual sample, compared to the expected frequency (Field 2013).

with relation to the number of references to the effector or to the target of the grasp is not significant ($\chi^2(6, N=322)=6.453, p>0.05$).

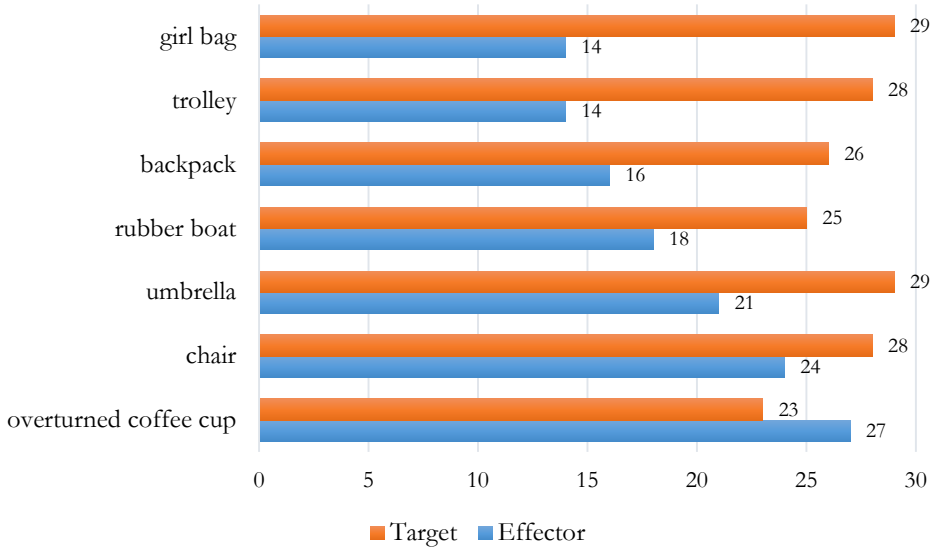



Figure 3.2. References to the effector (blue bars, tot. 134) and references to the target (red bars, tot. 188) found in grasp descriptions for artefacts with affording parts (presented without different orientation).

However, we see important differences with respect to the previous group of stimuli (cf. Fig. 3.1). Descriptions that contain references to the effector of the grasp are now generally less frequent and there is more variation among different stimuli (mean=19.14; SD=5.05). On the contrary, the target of the grasp is mentioned in most grasp descriptions (mean=26.86; SD=2.27). These results appear to be the opposite of those illustrated in the preceding paragraph.

| | | girl bag | trolley | backpack | rubber boat | umbrella | chair | overturned coffee cup | tot. |
|---|---|---|---|---|---|---|---|---|---|
| **Effector** | frequency | 14 | 14 | 16 | 18 | 21 | 24 | 27 | 134 |
| | std. residual | -0.9 | -0.8 | -0.4 | 0 | 0 | 0.5 | 1.4 | |
| **Target** | frequency | 29 | 28 | 26 | 25 | 29 | 28 | 23 | 188 |
| | std. residual | 0.8 | 0.7 | 0.3 | 0 | 0 | -0.4 | -1.1 | |

Table 3.3. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for artefacts with affording parts (presented without different orientation).

Looking in detail at Figure 3.2 and Table 3.3, we notice that within this category the overturned coffee cup is the visual stimulus that elicited the fewest mentions of the target (23) and the most mentions of the effector (27). It is also the stimulus for which the largest gap between the observed and the expected values (standardised residuals: 1.4 and -1.1) is recorded. These values are more similar to those presented for artefacts without affording parts than to those derived from other stimuli of this category: only in this case does the frequency of references to the effector exceed the frequency of references to the target. This is probably because the object presented is overturned and very small; for this reason, its handle or other graspable parts are not judged to be probable targets of the grasp and are thus rarely mentioned by informants (*manico*, "handle" occurs very rarely), who in most cases described a simple and undifferentiated power grasp, directed either to the sides of the object (*lato*, "side") or to the bottom of the cup (*fondo*, "bottom"; *alto*, "the upper part").

On the other hand, the girl bag and the umbrella are the stimuli for which there are the most descriptions (29 out of 30) that contain one or more references to the target of the grasp, followed by the trolley and the chair (28 occurrences).

### 3.2.3 Artefacts with affording parts (with different orientation)

For this second group of artefacts with affording parts (presented with different orientation), data from the rightward-oriented and leftward-oriented stimuli will first be merged, in order to discover if the objects used as stimuli (independently from their orientation) differ from each others with relation to the number of mentions of either the

target or the effector of the grasp described (Fig. 3.3). The two conditions will then be analysed separately (§3.2.3.1 and §3.2.3.2).

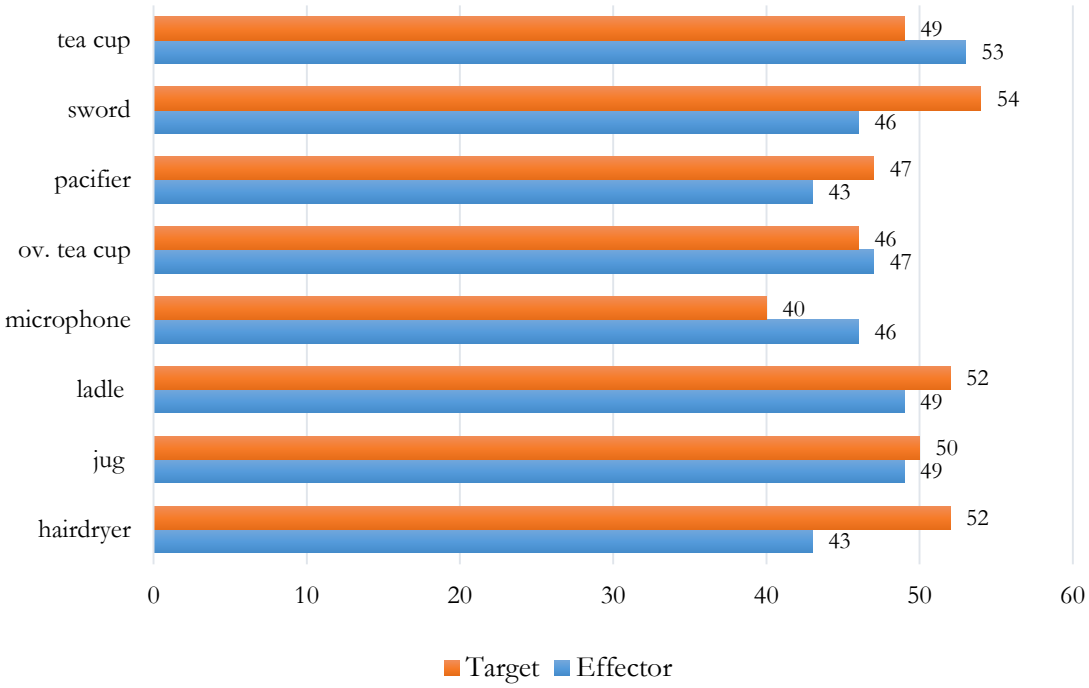



Figure 3.3. References to the effector (blue bars, tot. 376) and references to the target (red bars, tot. 390) found in grasp descriptions for artefacts with affording parts presented with different orientation. For each object, data from the rightward condition are merged with those for the leftward condition (tot. 60 descriptions per object).

Looking at Figure 3.3, we notice that, in general, descriptions that contain references to the effector of the grasp are slightly less frequent (mean=47; SD=3.12) with respect to those containing an explicit reference to the target of the grasp (mean=48.75; SD=4.15). These results are in line with those illustrated for the first group of artefacts with affording parts (cf. Fig. 3.2): in that case, too, the target was mentioned more frequently than the effector of the grasp.

As already observed for the first group of artefacts stimuli with affording parts, the differences between the objects used as stimuli with relation to the mention of either the target or the effector are ultimately not significant ($\chi^2(7, N=766)=2.101, p>0.05$; cf. also Tab. 3.4). In other words, different stimuli do not elicit a significantly different proportion of references to either the effector or the target of the grasp described.

| | | hairdryer | jug | ladle | microphone | overturned tea-cup | pacifier | sword | tea-cup | tot. |
|---|---|---|---|---|---|---|---|---|---|---|
| **Effector** | frequency | 43 | 49 | 49 | 46 | 47 | 43 | 46 | 53 | 376 |
| | std. residual | -0.5 | 0.1 | 0 | 0.6 | 0.2 | -0.2 | -0.4 | 0.4 | |
| **Target** | frequency | 52 | 50 | 52 | 40 | 46 | 47 | 54 | 49 | 390 |
| | std. residual | 0.5 | 0 | 0.1 | -0.6 | -0.2 | 0.2 | 0.4 | -0.4 | |

Table 3.4. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for artefacts with affording parts (presented with different orientation). For each object-stimulus, data from the rightward condition are merged with those for the leftward condition.

However, a more detailed analysis is required for this second category of artefacts with affording parts (eight objects, 16 stimuli), since there is another variable that has to be considered besides object typology: the different horizontal orientation of the objects-stimuli presented.

For this reason, we will now distinguish between the two groups of objects-stimuli (rightward-oriented vs. leftward-oriented), as well as between the descriptions provided by the group of right-handed informants and those provided by left-handed informants. Our purpose is to verify if there is an effect of object's orientation on lexical choices with relation to informants' hand dominance.

#### 3.2.3.1 Left-handed informants

We will first consider the descriptions provided by the left-handed informants (seven informants; tot. 112 descriptions). The following graphs represent the number of references to the target and to the effector (red and blue bars respectively), for the rightward orientation (first couple of bars in each graph) and the leftward orientation (second couple), contained in the seven descriptions provided for each object-stimulus.

The target of the grasp appears to be mentioned more frequently than (or at least as frequently as) the grasp effector. This is true for all stimuli except the microphone, albeit only in its leftward orientation. Considering all descriptions provided for the sixteen objects-

stimuli, we find 92 references to the target of the grasp (mean=5.75 per stimulus; SD=1.2) and 69 references to the effector (mean=4.31 per stimulus; SD=1).

However, a difference between the two different possible orientations in which stimuli are presented emerges.

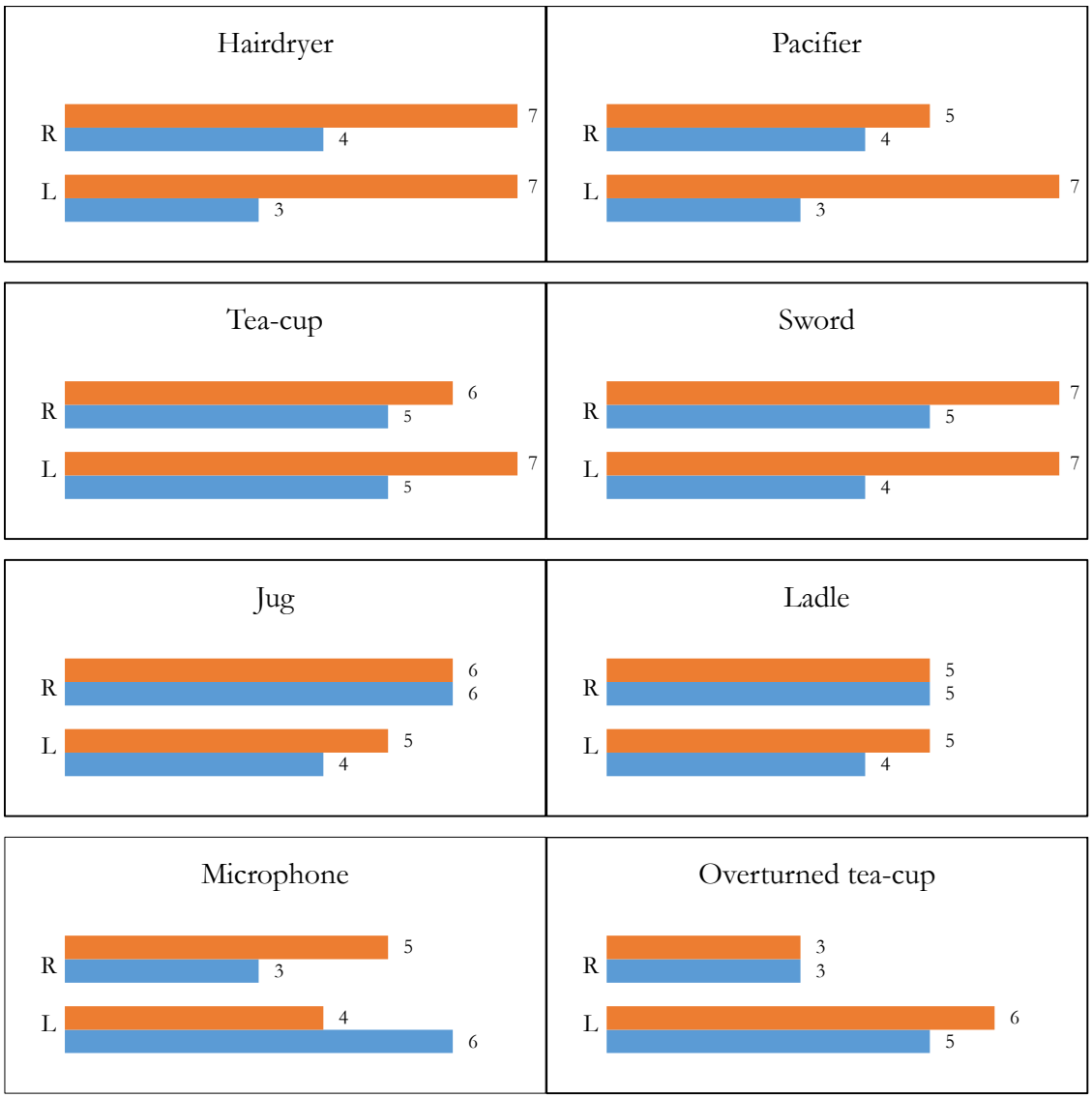



Figure 3.4. Data from the left-handed group (seven informants), for the artefacts presented with different orientation (eight objects; 16 stimuli). References to the target (red bars, tot. 92) and to the effector (blue bars; tot. 69) of the grasp elicited by either the rightward-oriented (first and second bars in each graph) or the leftward-oriented objects (third and fourth bar) are compared.

We notice a slight tendency to name the target of the grasp, usually the handle, especially when the object is leftward-oriented rather than in the opposite condition (the references to the target elicited by leftward-oriented objects are 48, whereas those elicited by rightward-oriented objects are 44). Only for two pairs of objects-stimuli, i.e. the microphone and the jug, is the target explicitated more often in the rightward condition than in the leftward condition.

On the other hand, references to the effector are slightly more frequent in the grasp descriptions provided for the rightward-oriented objects than for the leftward-oriented objects (35 vs. 34). Only in two cases, i.e. the microphone and the overturned tea-cup, are the references to the effector more frequent for the leftward condition than for the rightward condition.

Table 3.5 summarises the number of references to the target and to the effector of the grasp found in descriptions provided for rightward-oriented and leftward-oriented objects.

| | Target | Effector |
|---|---|---|
| Rightward-oriented | 44 | 35 |
| Leftward-oriented | 48 | 34 |

Table 3.5. Frequency of references to the target and to the effector of the grasp for the two groups of stimuli (the eight rightward/leftward-oriented artefacts) provided by the left-handed participants (seven informants; tot. 112 descriptions).

#### 3.2.3.2 Right-handed informants

We will now consider the descriptions provided by the right-handed informants for the stimuli already presented in the previous paragraph (23 informants; tot. 368 descriptions). Looking at the descriptions elicited for the 16 stimuli, references to the target are 298 (mean=18.63 per stimulus; SD=1.96), whereas references to the effector are 307 (mean=19.19 per stimulus; SD=1.7). This difference is very slight and, in many cases, the number of references to the effector exceeds the number of references to the target. However, this occurs mostly for leftward-oriented objects: the pacifier, the tea-cup, the jug, the ladle and the microphone. Only in two cases does the number of references to the

effector exceed the number of references to the target for rightward-oriented objects: the tea-cup, both upright and overturned.

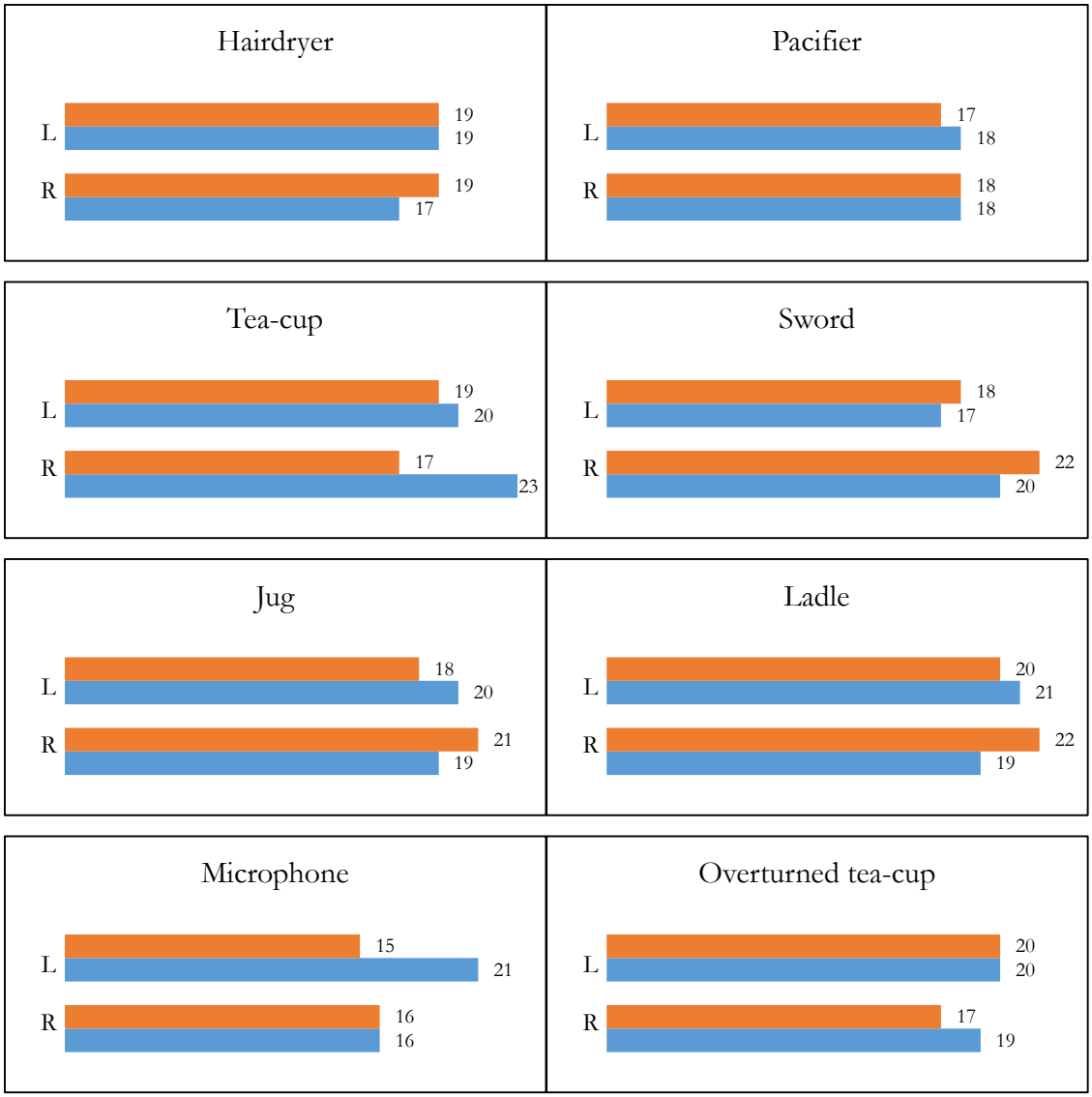



Figure 3.5. Data from the right-handed group (23 informants), for the artefacts presented with different orientation (eight objects; 16 stimuli). References to the target (red bars, tot. 298) and to the effector (blue bars; tot. 307) of the grasp elicited by either the leftward-oriented (first and second bars in each graph) or the rightward-oriented objects (third and fourth bar) are compared.

Therefore, as already noticed in the analysis for the left-handed group, in this case, too, there is a slight difference between the two orientation conditions (Tab. 3.6): references to the target appear to predominate in the rightward rather than in the leftward condition (152

vs. 146). On the contrary, descriptions that contain a mention of the effector are mostly found in relation to leftward-oriented (156) rather than to rightward-oriented objects (151).

| | Target | Effector |
|---|---|---|
| Rightward-oriented | 152 | 151 |
| Leftward-oriented | 146 | 156 |

Table 3.6. Frequency of references to the target and to the effector of the grasp for the two groups of stimuli (the eight rightward/leftward-oriented artefacts) provided by the right-handed participants (23 informants; tot. 368 descriptions).

#### 3.2.3.3 Comparing the two groups of informants: the effects of object orientation and hand dominance

In the descriptions provided by both groups of informants, we observed a tendency to name the target of the grasp especially when it is spatially aligned with the dominant hand. Is it possible to assert that the spatial alignment between object orientation and hand dominance really influences lexical choices in grasp descriptions? To answer this question, we can now merge data from Tab. 3.5 and Tab. 3.6 into Tab. 3.7 in which, instead of object orientation (rightward vs. leftward) and hand dominance (right-handed vs. left-handed), we consider whether the orientation of the affording part is spatially aligned with the dominant hand of the informants or not.

| | Target | Effector |
|---|---|---|
| + Spatial alignment | 200 | 185 |
| - Spatial alignment | 190 | 191 |

Table 3.7. References to the target and to the effector of the grasp in two conditions, i.e. of spatial alignment or non-spatial alignment between the orientation of the 16 object-stimulus and the hand dominance of the 30 informants (480 descriptions).

It is clear that spatial alignment has only a very slight (and not significant) effect on the references to the target or the effector of the grasp ($\chi^2(1, N=766)=0.331$, $p>0.05$). The probability of finding an explicit mention of the target or the effector of the grasp, for this

group of stimuli, appears to critically depend on neither the orientation in which the object is presented, nor the hand dominance of the informant.

However, this result was to be expected. The initial hypothesis that guided the linguistic analysis of the transcripts was that the frequency of target-related words may reflect the salience of objects' affording parts. Data suggest that there is only a slight difference between the descriptions provided for stimuli that are spatially aligned with the dominant hand of informants and the descriptions provided for those that are not spatially aligned, for what regards the number of references to the target of the grasp (that in most cases, for these objects, denote affording parts; cf. §4.3.3.1). This suggests that the salience of the affording parts does not depend on a mere spatial alignment effect between object orientation and hand dominance. The handle always maintains its strong affording power, regardless of the object's orientation. Nevertheless, spatial alignment turns out to have a small facilitation effect: in the grasp descriptions collected from both groups of informants, explicit mentions of either the target or the effector are more frequent for objects-stimuli that are, respectively, spatially aligned and non-spatially aligned with the dominant hand of the informant (Tab. 3.7).[28]

We can now turn to compare the two categories of artefacts without affording parts and artefacts with affording parts.

### 3.2.4 General remarks on the artefacts category: the effect of affording parts on linguistic production

The main difference between the two groups of artefacts objects-stimuli (with or without affording parts) regards the information about the effector or the target of the grasps that can be retrieved in the descriptions provided by informants.

First of all, grasps directed to artefacts with affording part (e.g. the cup) are generally described by explicitating the effector or the target of the action (e.g. the hand, the handle) more frequently than grasps directed to artefacts without affording parts (e.g. the ball).

[28] The effect of objects' orientation on the type of grasp described by informants, in particular with respect to the interference of hand dominance on the spatial compatibility effect, will be further investigated in a case study presented in §6.2.

As already stated, since we have 30 grasp descriptions for each object, the total number of (effector/target) references collected for each stimulus may range, in theory, from 0 to 60: 0 in the very unlikely case that none of the thirty informants named the effector or the target of the grasp; 60 in the case that all the 30 informants mentioned both the effector and the target. Notably, the artefacts stimuli that are characterised by the lowest number of references (to either the target or to the effector) are all objects without affording parts, namely, the pencil (33 effector/target references), the lighter (35 references), the tennis ball, and the soccer ball (37 references). For these objects-stimuli, most informants provided very short descriptions, with only one reference to either the effector or the target of the grasp. The eight artefacts without affording parts have indeed a mean of 40.75 references for each stimulus (SD=6.23; references per informant per object: 1.36).

On the contrary, the artefacts stimuli that are characterised by the highest number of references are the rightward-oriented sword (54 references), the chair, the rightward-oriented jug (52 references each), the rightward-oriented ladle, the upright tea-cup (in both orientation), and the overturned leftward-oriented tea-cup (all with 51 effector/target references). For these objects, most descriptions contain the reference to both the effector and the target of the grasp. The 23 artefacts with affording parts have a mean of 47.3 references for each stimulus (SD=4.02; references per informant per object: 1.58).

Therefore, grasp descriptions provided for artefacts with affording parts usually contain a higher number of references to either the effector or the target of the action, with respect to the descriptions provided for artefacts without affording parts. However, a more interesting difference can be observed between the two groups. The descriptions provided for the artefacts without affording parts contain in proportion more references to the effector (66%) than references to the target of the grasp (34%), whereas, on the contrary, descriptions provided for the artefacts with affording parts contain more references to the target (53%) than to the effector (47%) of the grasp (Fig. 3.6).

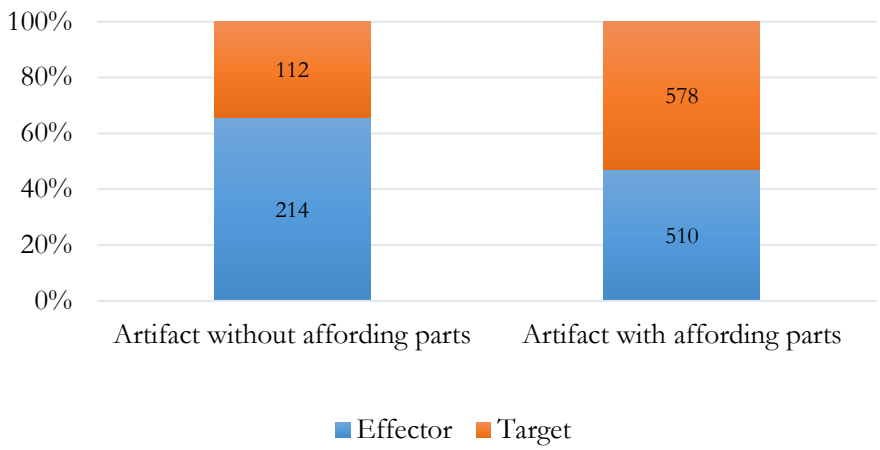



Figure 3.6. Percentages of references to the effector and the target of the grasp for artefacts with affording parts and artefacts without affording parts. Frequency data are reported in labels.

There is an evident difference between the descriptions provided for the two kinds of stimuli (artefacts with and without affording parts), with relation to the number of references to the effector or the target of the grasp ($\chi^2(1, N=1414)=35.367$, $p<0.001$). Descriptions provided for the eight artefacts without affording parts are characterised by a higher number of references to the effector (mean=26.75; SD=2.19) and a lower number of references to the target of the grasp (mean=14; SD=7.6), with respect to those provided for 23 artefacts with affording parts, in which the number of references to the target is much higher (mean=25.13; SD=2.8), and even exceeds the number of references to the effector (mean=22.17; SD=3.82).

Such results seem to indicate that the salience of the affording part is reflected in the linguistic production. When artefacts present affording parts, most informants explicitly name the target of the grasp. Moreover, the more the target of the grasp is named, the less the effector of the grasp is mentioned. Therefore, linguistic elements can be seen as cues for a shift of attention towards the object's parts that allow and facilitate the grasp. On the contrary, when artefacts do not present affording parts, subjects typically name their hand or their fingers, but it is much less frequent that they mention the target of the action.

## 3.3 Humans

We can now turn to the linguistic analysis of the 30 descriptions provided for the stimuli of the human class. Figure 3.7 reports the number of references to the effector and to the target of the grasp contained in the descriptions provided for different objects-stimuli, i.e. the baby, the standing girl, and the running man (tot. 90 descriptions).

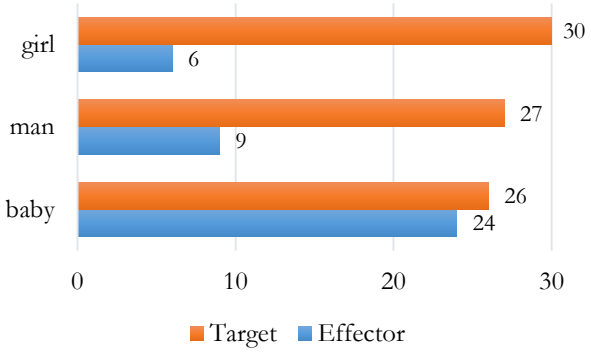



Figure 3.7. References to the effector (blue bars, tot. 39) and references to the target (red bars, tot. 83) found in grasp descriptions for human kinds.

In the descriptions provided for the three human kinds, the number of references to the target of the grasp is very high (mean=27.67; SD=1.7), with respect to the number of references to the effector (mean=13; SD=7.87). However, there are evident differences among the descriptions provided for these three stimuli ($\chi^2(2, N=122)=10.588$, $p<0.01$). In particular (cf. Tab. 3.8), we observe that the image of the baby stimulates a particularly high number of mentions of the effector of the grasp (that is the main responsible for the outcome of the chi-square test). This is because most participants specified that they would have grasped the baby with two hands and described a very delicate and careful grasp (cf. §4.4.1). For the girl and the man, words related to the effector are much rarer in the transcripts.

On the contrary, in all descriptions collected for this group of stimuli there is a high number of mentions of the target of the grasp, that even exceeds that found for the category of artefacts with affording parts (as will be discussed in detail in §4.4.1, the target is mostly indicated by body parts, such as *braccio*, "arm", *gomito*, "elbow", *ascelle*, "armpits").

| | | girl | man | baby | tot. |
|---|---|---|---|---|---|
| **Effector** | frequency | 6 | 9 | 24 | 39 |
| | std. residual | -1.6 | -.7 | 2 | |
| **Target** | frequency | 30 | 27 | 26 | 83 |
| | std. residual | 1.1 | .5 | -1.4 | |

Table 3.8. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for human kinds.

## 3.4 Natural kinds

There are four natural kinds among the stimuli adopted in the experiment: they are the stone, the mandarin, the banana, and the apple. For these objects-stimuli, informants provided a total amount of 120 grasp descriptions that contain 148 references to either the effector or the target of the grasp (Fig. 3.8).

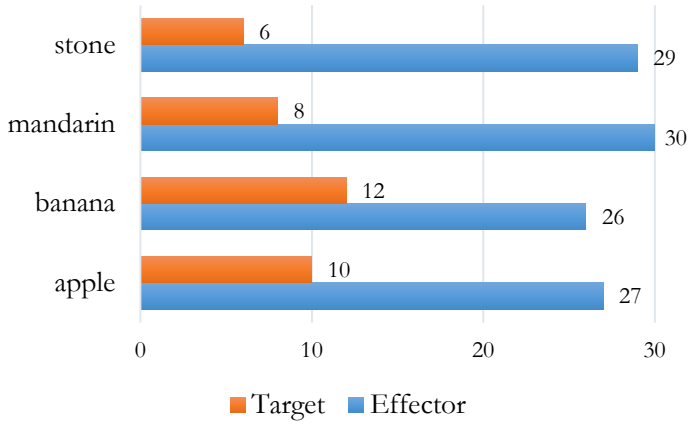



Figure 3.8. References to the effector (blue bars, tot. 112) and references to the target (red bars, tot. 36) found in grasp descriptions for human kinds.

There is no significant difference between the four stimuli (cf. also Tab. 3.9), with respect to the number of references to either the effector or the target of the grasp found in transcripts ($\chi^2$ (3, $N$=148)=2.435, $p$>0.05). Almost all descriptions collected for each stimulus contain an explicit mention of the effector of the grasp (mean=28 per stimulus; SD=1.58). On the other hand, mentions to the target are much rarer, ranging from six to 12 occurrences (mean=9 per stimulus; SD=2.24).

| | | apple | banana | mandarin | stone | tot. |
|---|---|---|---|---|---|---|
| **Effector** | frequency | 27 | 26 | 30 | 29 | 112 |
| | std. residual | -1 | -2.8 | 1.2 | 2.5 | |
| **Target** | frequency | 10 | 12 | 8 | 6 | 36 |
| | std. residual | 1 | 2.8 | -1.2 | -2.5 | |

Table 3.9. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for natural kinds.

In general, this category is characterised by a higher number of references to the effector to the grasp, and less frequent mentions to the target, than any other category of stimuli analysed so far.

## 3.5 Substances and aggregates

Within this group of objects-stimuli there are four elements: the water, the sand, the pumpkin seeds, and the flour.

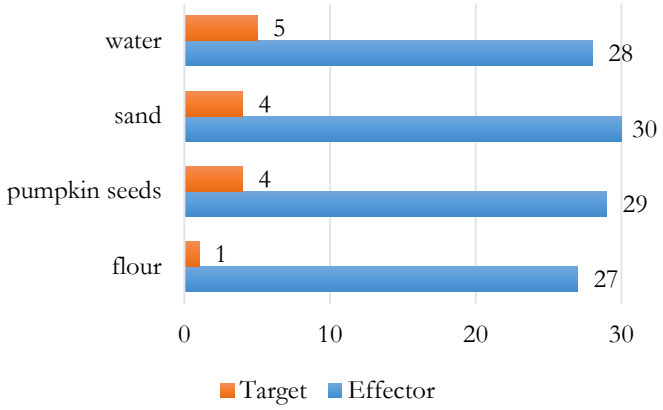



Figure 3.9. References to the effector (blue bars, tot. 114) and references to the target (red bars, tot. 14) found in grasp descriptions for substances and aggregates.

As Figure 3.9 shows, this class of stimuli is characterised by a very high number of references to the effector of the grasp (mean=28.5 per stimulus; SD=1.12) that exceeds that observed for other categories analysed so far (even natural kinds). On the other hand, mentions to the target of the grasp are rarer, compared to those of other stimuli (mean=3.5;

SD=1.5). In fact, they are the rarest of all stimuli: only for the pencil we observed such a low frequency of references to the target of the grasp (five references, as for the water).

Such a low number of references to the target of the grasp is certainly due to the difficulty in recognising and naming the specific parts of these entities that could constitute the target of the grasp, i.e. the exact place where the contact with the effector is intended to occur. This topic will be better explored in the next chapter (cf. §4.6.2; §4.7.1.3; §4.7.2.2).

## 3.6 General results

We can now compare the results obtained from the analysis of the different classes of stimuli and draw some conclusions. Table 3.10 gathers the data already presented for each stimulus.

| Object-stimulus | Effector | Target |
|---|---|---|
| apple | 27 | 10 |
| baby | 24 | 26 |
| backpack | 16 | 26 |
| banana | 26 | 12 |
| box | 28 | 22 |
| chair | 24 | 28 |
| coffee cup O | 27 | 23 |
| flour | 27 | 1 |
| girl | 6 | 30 |
| girl bag | 14 | 29 |
| glass | 28 | 14 |
| hairdryer L | 22 | 26 |
| hairdryer R | 21 | 26 |
| jug L | 24 | 23 |

| Object-stimulus | Effector | Target |
|---|---|---|
| jug R | 25 | 27 |
| ladle L | 25 | 25 |
| ladle R | 24 | 27 |
| lighter | 28 | 7 |
| man | 9 | 27 |
| mandarin | 30 | 8 |
| microphone L | 27 | 19 |
| microphone R | 19 | 21 |
| pacifier L | 21 | 24 |
| pacifier R | 22 | 23 |
| pencil | 28 | 5 |
| plate | 25 | 19 |
| pumpkin seeds | 29 | 4 |
| rubber boat | 18 | 25 |

| Object-stimulus | Effector | Target |
|---|---|---|
| sand | 30 | 4 |
| soccer ball | 27 | 10 |
| stone | 29 | 6 |
| sword L | 21 | 25 |
| sword R | 25 | 29 |
| tea-cup L | 25 | 26 |
| tea-cup R | 28 | 23 |
| tea-cup O, L | 25 | 26 |
| tea-cup O, R | 22 | 20 |
| tennis ball | 28 | 9 |
| trolley | 14 | 28 |
| umbrella | 21 | 29 |
| vase | 22 | 26 |
| water | 28 | 5 |

Table 3.10. References to the target and to the effector found in grasp descriptions provided for 42 objects-stimuli.

As it can be observed from Table 3.10, the references to the effector are quite frequent for all objects-stimuli (tot. 989; mean=23.55; median=25; mode=28), even if there is an

evident variation among the values (SD=5.36), which range from six to 30. In particular, only four stimuli did not elicit any explicit mention of the effector of the grasp in at least the 50% of the descriptions (references to the grasp effector <15), whereas for 11 stimuli are the references to the effector equal to or greater than 28. The mentions of the target of the grasp are instead less frequent (tot. 823; mean=19.6; median=23; mode=26) and they show even more variation (SD=8.8), ranging from one to 30. The objects-stimuli that did not elicit any explicit reference to the target of the grasp in at least 15 descriptions are now 13, whereas for only six stimuli are the references to the target equal to or greater than 28.

We can now observe the correspondence between the mention of the target/effector of the grasp and the kind of object-stimulus presented. As it appears evident from Figure 3.10, some kind of correspondence does exist.

In Figure 3.10, the different objects-stimuli are coloured in different shades of blue, according to their semantic type. We use, here, the term "semantic type", even if we are actually also distinguishing, within the artefact category, between artefacts with and without affording parts. Frequency values are coloured in shades of red, in order to indicate which stimuli have the lowest (yellow) or highest (red) number of references to either the effector or the target of the grasp. Objects-stimuli are presented in two separated lists: on the left, they are ordered according to the number of references to the effector of the grasp, whereas, on the right, they are ordered according to the number of references to the target.

Focussing on the left side of the figure, it is possible to notice an evident gathering of cells coloured in light blue in the bottom part (natural kinds, substances and aggregates, characterised by a very high number of references to the effector of the grasp), whereas in the same column most of the dark blue cells (human kinds and artefacts with affording parts, characterised by a low number of references to the effector) occupy an upper position. The right side of the figure, related to the number of references to the target of the grasp, shows the opposite situation: the light blue cells (natural kinds, substances and aggregates) are concentrated in the upper part, whereas at the bottom we find mostly cells coloured in dark blue (human kinds and artefacts with affording parts). In their middle part, the two lateral columns are quite similar in colour.

The two central columns, besides containing frequency data with relation to the lateral columns, also show what we have already noticed about the distribution of the references found in transcripts: mentions of the effector are, in general, more frequent than mentions of the target (the left side is darker).

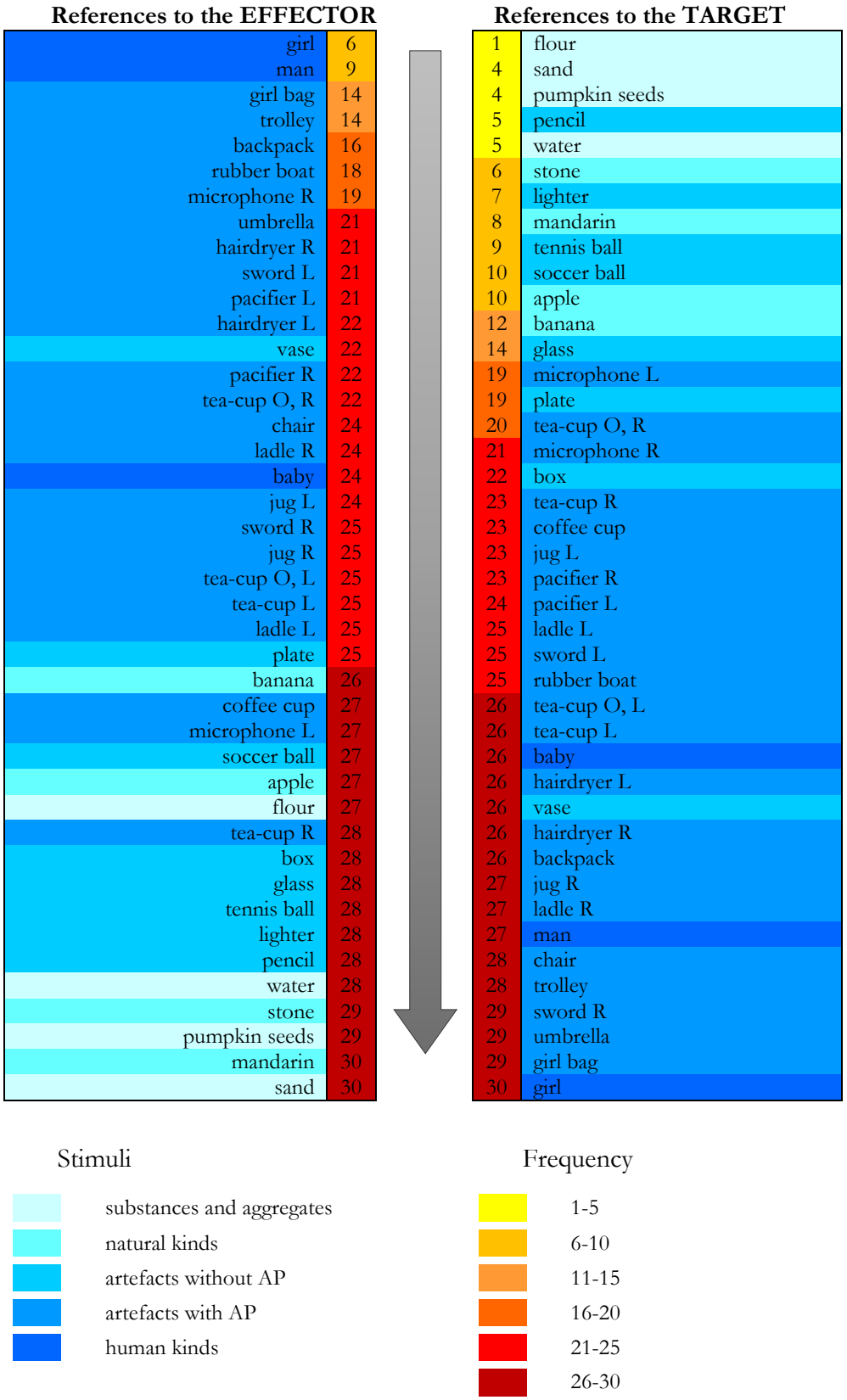



Figure 3.10. References to the target and references to the effector found in grasp descriptions provided for each object-stimulus.

To verify whether the impression got from this figure is correct, we can now merge data from the different objects-stimuli into five categories. First of all, we consider absolute frequencies (Table 3.11):

| | N° stimuli | References to the Effector | References to the Target | Tot. references |
|---|---|---|---|---|
| Human kinds | 3 | 39 | 83 | 122 |
| Artefacts (with AP) | 23 | 510 | 578 | 1088 |
| Artefacts (without AP) | 8 | 214 | 112 | 326 |
| Natural kinds | 4 | 112 | 36 | 148 |
| Substances and Aggregates | 4 | 114 | 14 | 128 |
| Tot. | 42 | 989 | 823 | 1812 |

Table 3.11. Number of references to the target and to the effector found in grasp descriptions provided for the five classes of stimuli (absolute frequencies; AP= affording parts).

While for the class of human kinds the references to the target are twice as frequent as the references to the effector, for artefacts with affording parts the two values are much closer. As regards artefacts without affording parts, natural kinds, and substances and aggregates, the number of references to the effector strongly increases over that of references to the target.

Therefore, there are evident and highly significant differences between the five types of stimuli for which grasp descriptions have been provided by informants, with relation to the number of references either to the effector or to the target of the grasp ($\chi^2(4, N=1812)=155.3, p<0.001$; cf. Tab. 3.12).

| Class of stimuli | | Effector | Target | Tot. |
|---|---|---|---|---|
| Human | frequency | 39 | 83 | 122 |
| | std. residual | -3.4 | 3.7 | |
| Artefacts with AP | frequency | 510 | 578 | 1088 |
| | std. residual | -3.4 | 3.8 | |
| Artefacts without AP | frequency | 214 | 112 | 326 |
| | std. residual | 2.7 | -3 | |
| Natural kinds | frequency | 112 | 36 | 148 |
| | std. residual | 3.5 | -3.8 | |
| Substances and aggregates | frequency | 114 | 14 | 128 |
| | std. residual | 5.3 | -5.8 | |
| | freq. tot. | 989 | 823 | 1812 |

Table 3.12. Crosstabulation and standardised residuals of the references to the effector or to the target found in the grasp descriptions provided for the five categories of stimuli.

Finally, since the classes contain a very different number of stimuli, it is also useful to look at percentages. The following graph (Fig. 3.11) shows the percentages of references to the target and to the effector of the grasp over the total amount of references collected for each class.

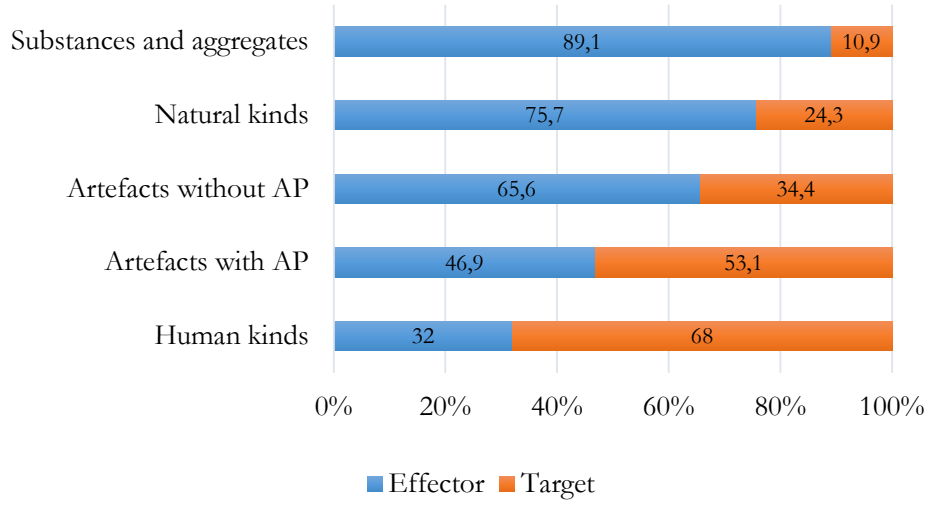



Figure 3.11. References to the target and to the effector found in grasp descriptions provided for the five classes of stimuli (percentages over the total number of references extracted).

This 100% stacked bar graph better illustrates what we have noticed above: that substances and aggregates are characterised by a very high number of references to the effector of the grasp, compared to references to the target. For natural kinds and artefacts without affording parts, the number of references to the target increases, although most mentions are still referred to the effector. With the artefacts with affording parts, and especially with humans, the percentage of mentions to the target exceeds that of mentions to the effector of the grasp.

## 3.7 Discussion

Considering the number of references to the effector or to the target of the grasp, we can create a hierarchy such as the following:

human kinds > artefacts with AP > artefacts without AP > natural kinds > substances/aggregates

The more the object-stimulus is on the left part of this hierarchy, the more likely the target of the grasp is to be named in the descriptions, and the less likely the effector of the grasp is to be mentioned.

This hierarchy well describes the data presented in the previous paragraphs. We have observed that the mean frequency of references to the effector increases, moving from the left to the right side of the hierarchy (humans: mean=13; artefacts with affording parts: mean=22.17; artefacts without affording parts: mean=26.75; natural kinds: mean=28; substances and aggregates: mean=28.5). Parallel to this, the frequency of references to the target decreases (humans: mean=27.67; artefacts with affording parts; mean=25.13; artefacts without affording parts: mean=14; natural kinds: mean=9; substances and aggregates: mean=3.5).

Therefore, the descriptions of possible grasps provided by informants after the presentation of a natural kind or a substance/aggregate stimulus, are characterised by a predominance of references to the effector of the grasp. On the contrary, the descriptions of possible grasps directed towards humans and artefacts with affording parts are characterised by a higher number of mentions of objects' parts that constitute the target of the action.

The results of this first study conducted on the grasp descriptions suggest that informants' answers reflect shifts of attention from the agent-side to the object-side with relation to different categories of stimuli. We can now tackle a more in-depth analysis of the transcripts, aimed at investigating the lexical expressions used by informants in denoting the effector and the target of the grasp. A broader and more comprehensive discussion will be provided at the end of the next chapter.

# Chapter IV

# The Language of Affordances

In Chapter III, explicit references to either the target or the effector of the grasp found in the descriptions were analysed; now, the linguistic content of the transcripts can be examined in more detail. The main purpose of this study is to highlight which are the lexical expressions used by informants to denote the effector and the target of the grasp.

There are two main methodological differences between the analysis presented in this chapter and the one described in Chapter III.

The first one is that while in the previous analysis each description was only supposed to contain a maximum of one reference to the effector and one reference to the target of the grasp, in this more in-depth study *all* lexical expressions extracted from answers will be reported, in order to highlight how the linguistic reference to the effector and the target of the grasp is made.[29] This will allow us to discover other differences among objects' types. For example, a description such as «la scatola la prenderei mettendo le mani ai lati e sorreggendolo dalla base, quindi il palmo andrebbe a toccare i due lati e le dita andrebbero a

[29] Repetitions of the very same word within the same description, as frequently happens in speech, are counted as a single occurrence.

toccare il sotto, la base, in modo da poterlo sollevare»,[30] besides being very detailed and informative, is interesting from a linguistic point of view because the reference to the target is made through the reference to its sides, its bottom, and its base. These are all lexical expressions pertaining to the domain of spatial relations: the base indicates the lowest part of the object; the sides are the surfaces of an object that are not in the front, back, top, or bottom and here probably are the two faces of the box aligned with the informants' hands; *bottom* is also a word clearly pertaining to the spatial domain.

On the contrary, a frequent description provided for the rightward-oriented cup is «con una mano, dal manico», with a reference to a single hand and, moreover, with a reference to a specific part of the object (here, the handle). This is not because this specific content could not be expressed in terms of spatial relations or be left understood: in fact, in place of "from the handle", the informant could have also said "from its right side" (maintaining the reference to the target), or just "with my right hand" (referring only to the effector).

The second significant difference is that this second study also aims to go beyond and broaden the simple concepts of *reference to the effector* and *reference to the target* on which the previous analysis was based, where effector and target were defined as follows (§3.1.1):

- the *effector* of the grasp is the entity that is linguistically presented as the one that comes in contact with the object;
- the *target* of the grasp is the part of the object-stimulus where the contact with the effector is described to occur.

In this chapter, certain kinds of lexical expressions that are *related* to the effector of the grasp or to its target are also taken into consideration, even if they do not indicate the effector or the target in a direct and precise form, but qualify or quantify it (the differences from the previous analysis will be better specified in what follows). For instance, the meronym *dita*, "fingers", perfectly specifies the part of the hand with which a grasp may be performed, and therefore constitutes a true reference to the effector; however, a grasp described as being executed with the fingers is different from one performed with *two* fingers or with *all* of the

[30] Since this chapter will be rich in examples taken from transcripts, they will be incorporated into the text, as quotations of the informants' speech enclosed within double angle brackets. In order to streamline the reading, the notations and symbols described in Chapter II (cf. in particular §2.4.5), as well as retractions and repetitions, are not reported.

fingers: such expressions may be used to describe a precision or a power grip, respectively. This very basic example shows that, even if in most cases expressions of quantity alone do not suffice to explicitate neither the effector nor the target of the grasp, they are indeed well worth considering in this more detailed analysis.

In order to conduct the qualitative and quantitative analysis of the words used to indicate either the effector or the target of the grasp, all target-related and effector-related words were classified adopting a set of semantic categories. Such method is inspired by the research conducted on feature norms, which are semantic properties of concepts (usually evoked by the written form of a word) collected empirically with a property generation task.[31] All different schemas proposed to classify semantic feature norms (e.g. McRae et al. 2005; Vinson and Vigliocco 2008; Wu and Barsalou 2009; Kremer and Baroni 2011; Montefinese et al 2013; Lenci et al. 2013) share the idea that the properties generated by informants may be classified according to the semantic relation that they establish with a given stimulus. In the present work, too, the words referred to the effector or the target of the grasp are classified according to the semantic relation that they establish with the entity described as the effector or with the object-stimulus presented during the task (for instance, the word *picciolo*, "stalk", denotes a part of the apple and is classified as a meronym). However, this classification, although being to some extent inspired by the one proposed in Lenci et al. 2013 (which will also be adopted to classify the feature norms collected in the experiment described in Chapter V; cf. in particular §5.2.4) presents many peculiarities and cannot be strictly compared to any existing classification. This is mostly due to the specificity of the experiment conducted, which is not a property generation task, but rather an action description task. In particular, the only stimuli adopted are images representing graspable entities (cf. §2.4.3): therefore, when informants describe the effector of the grasp (e.g. the hand), they are actually *introducing* an element that is not provided as stimulus, neither is a property of the visual stimulus itself (e.g. a cup), but it is present in the event of grasping that participants are requested to describe.

[31] This topic will be addressed in more detail in Chapter V (cf. in particular §5.1).

## 4.1 The classification of effector-related words

The semantic categories, into which all words related to the effector extracted from the transcripts were gathered, in most cases pertain to the domain of a manual grasp (categories 1-7). However, in very few cases, informants referred to a grasp performed in a non-canonical form, either with another body part, such as the mouth (8), or with an instrument[32] (9):

1) **Hand** *mano* ("hand"), *mani* ("hands")
   e.g. prendere la mela con la *mano*
   2) **Meronym** *dita* ("fingers"), *palmo* ("palm"), *pollice* ("thumb")
   e.g. prendere la mela con le *dita*
   3) **Holonym** *braccio* ("arm"), *braccia* ("arms")
   e.g. prendere la scatola con tutte le *braccia*
   4) **Space** *destra* ("right"); *sinistra* ("left")
   e.g. prendere la mela con la mano *destra*
   5) **Perceptive** *concavo* ("concave")
   e.g. prendere l'acqua con la mano *concava*
   6) **Quantity** *due* ("two); *tutto* ("whole"/"all")
   e.g. prendere la scatola con *due* mani
   7) **Similes** (based on perceptual properties) *coppa* ("cup"); *contenitore* ("container")
   e.g. prendere la sabbia con le mani a forma di *coppa*
8) **Other body part** *piede* ("foot"), *bocca* ("mouth")
   e.g. prendere il ciuccio con la *bocca*
9) **Instrument** *bicchiere* ("glass"), *tazza* ("cup")
   e.g. prendere l'acqua col *bicchiere*

Each semantic class is now presented in detail.

[32] Expressions pertaining to the semantic classes listed from (2) to (7) could, in theory, be related also to another body part or to an instrument; however, in the transcripts, they occur only in the description of a manual grasp.

1) **Hand (hnd)** In the linguistic descriptions of grasp collected in this study, the word *mano*, "hand", is the most frequent word related to the sphere of the effector (for frequency data relative to effector-related words, cf. §4.7.1). Not the fingers, nor the arms, but simply the hand. Therefore, frequency data extracted from transcripts strongly suggest that in the awareness of informants the basic (we would be tempted to say prototypical) effector of the grasp is the hand.[33] The first class of effector-related words is thus constituted by the lemma *mano*, together with *pugno*, "fist", which also refers to a hand, but shaped in a particular way (this is the only word that not only denotes the hand, but also expresses a modality). This also allows us to consider the hand as the central node of a chain, arm > hand > fingers, in which a hand has fingers and is part of an arm (cf. Cruse 1986: 160).

2) **Meronym (mer)** Considering the reference to the hand as the basic indication of the effector (mostly relying on frequency data), other effector-related words were classified in terms of the relation that they establish with the word *mano* and all of the words referring to the parts of the hand (such as fingers, the palm) were gathered in the category of (hand) meronyms. Expressions denoting the fingers are the most frequently attested within this class, and they can be considered as canonical meronyms of the hand (Cruse: 1986: 162).

3) **Holonym (hol)** Holonyms are words referring to the whole of which the hand is a part (typically *braccio*, "arm"). The arm is here considered as a holonym of the hand, and not as a separate effector (i.e. another body part), because the reference to the arms always implies that the hands are involved in the action together with the arms, as the detailed qualitative analysis provided in the next paragraphs will show. For instance, when subjects describe a grasp of a large box and refer to arms, it is implicitly understood, and often explicitly stated, that the hands hold the objects and the arms help to sustain it. Therefore, the grasp performed with two arms is always a grasp performed with the hands and *also* the rest of the arm, as a sort of extension of the dimension of the effector required by an object's size.

4) **Space (spa)** A single category collects all words denoting spatial relations referred to the effector of the grasp. Since the hands are two different effectors opposed to one

[33] Cf. §4.7.1.1 for a more detailed discussion.

another on a lateral axis, participants mostly adopted spatial terms to refer to the side of the effector considered, i.e. (*mano*) *destra* or *sinistra* ("right" or "left" hand); the left-right orientation of the hands is relative to the body of the informant, and can therefore be defined as egocentric (cf. Meini 2010: 23-25). Other kinds of words denoting spatial concepts with reference to the effector are very rare and will be discussed during the analysis.

5) **Perceptive (per)** This class gathers adjectives that express perceptive properties of the effector; for example, with relation to the hand's shape, such as *mano concava* ("concave hand").

6) **Quantity (qua)** This category is for number words, quantifiers, and all words expressing quantitative properties or attributes of the effector (or effector's parts) involved in the grasp. Sometimes, informants felt the need to specify that the "whole hand" is involved in the action (*mano intera*), or that the grasp of a pencil is performed by "two fingers" (*due dita*), or that an apple can be grasped with "only one hand" (*una mano sola*).[34]

7) **Similes, based on perceptual properties (spp)** This word class collects all nouns denoting a concrete entity evoked by informants in order to better describe a temporary property of the effector of the grasp, usually a particular hand shape. The nouns collected in this category are in most cases presented by informants as similes and analogies (e.g. *con la mano come se fosse un cucchiaio*, "with the hand as if it were a spoon").

8) **Other body part (obp)** Finally, for the very few cases in which a different body part (e.g. the foot or the mouth) is described as being the effector of the grasp, the category *other body part* was introduced.

9) **Instrument (ins)** In a few cases, the grasp described is performed by an instrument (e.g. a cup, a container) and the *instrument* category is adopted.

[34] We did not consider the occurrences of the lemma *uno*, "a/one" (e.g. *una mano*) as expressions of quantity (since in Italian it functions also as an indefinite article), but only the adjectives or adverbs that emphasise it (*una mano sola*, *soltanto una mano*, *un'unica mano*).

## 4.2 The classification of target-related words

The semantic categories into which all of the words referred to the target of the grasp were gathered are the following:

1) **Entity** *mela* ("apple"), *tazza* ("cup")
   e.g. prendere la *mela*
2) **Meronyms** *picciolo* ("stalk"), *manico* ("handle")
   e.g. prendere la mela dal *picciolo*
3) **Space** *sopra* ("above"), *lateralmente* ("laterally")
   e.g. prendere lo scatolone *lateralmente*
4) **Quantity** *due* ("two"), *tutto* ("all")
   e.g. prendere le *due* gambe della sedia
5) **Perceptive** *rosso* ("red"), *tondo* ("rounded")
   e.g. prendere la parte *rossa* del ciuccio
6) **Associated entity** *borsa* ("bag"), *sciarpa* ("scarf")
   e.g. prendere la donna per la *sciarpa*
7) **Similes** (based on perceptual properties) *pallina* ("small ball"), *sabbia* ("sand")
   e.g. prendere la farina come la *sabbia*

Again, each of these semantic classes will be briefly discussed in turn.

1) **Entity (ent)** Whereas, in the previous chapter, we took in consideration only the references to the target of the grasp that are effective in indicating the precise point of the object towards which the effector is directed (this point usually being expressed either by meronyms or by spatial expressions, for which see *infra*), the analysis is now also extended to some linguistic expressions that are still related to the object-stimulus. For example, when a participant, in front of a picture of a cup, says: «I would grasp this cup», he is not providing an informative answer, because he is simply repeating the information already provided in the initial instruction ("describe in the most detailed way how you would grasp this object") and naming the object-stimulus. However, at this finer-grained analysis, also the words denoting the object as a whole were extracted from the descriptions and classified, when they are clearly related to (or even presented as) the target of the grasp (e.g. *prendere la tazza*, *mettere*

*una mano intorno alla tazza*). Synonyms and hypernyms were also extracted, together with objects' names (e.g. *oggetto*, *strumento*, "object", "instrument").

2) **Meronym (mer)** This category gathers all expressions related to objects' parts that are presented as constituting the target of the grasp, such as the stalk of a fruit, or the handle of an umbrella.

3) **Space (spa)** The reference to the target of the grasp is often made through lexical expressions that pertain to the domain of space. We gather here, into a single, broad category, a very rich set of words that are used to denote spatial notions:

a. Nouns: such as *lato*, "side"; *base*, "base"; *centro*, "centre"
   e.g. prendere entrambi i *lati* dello scatolone; prendere la matita dal *centro*
b. Adjectives: such as *inferiore*, "lower"; *destro*, "right"; *sinistro*, "left"
   e.g. prendere lo scatolone dalla base *inferiore*
c. Adverbs: such as *lateralmente*, "laterally", *frontalmente*, "frontally"
   e.g. prendere lo scatolone *lateralmente*
d. Prepositions: such as *dietro*, "behind", *sotto*, "beneath"
   e.g. prendere lo scatolone con una mano *sotto* all'oggetto

For what regards prepositions, some further remarks have to be added. Italian has many locative prepositions, but according to morphosyntactic and semantic criteria, only the polysyllabic, secondary adpositions expressing spatial relations were considered (according to the Italian grammatical tradition, the so-called *preposizioni improprie*, "improper prepositions", such as *sotto*, *sopra*, *davanti*, *dietro*, *intorno*). Secondary prepositions[35] usually govern their complement by the intervention of certain other prepositions; however, most of them are polyfunctional, and the same words may function as adverbs (e.g. prepositional use: *stare dietro la porta*, "to stand behind the door", vs. adverbial use: *stare dietro*, "to stand

[35] Rizzi (1988: 521-522) distinguishes three classes, among Italian polysyllabic prepositions: I) those that mandatorily require a monosyllabic preposition (e.g. *accanto a*, "beside"); II) those that may admit a monosyllabic preposition, which is always *a* (e.g. *sopra*, "above", *sotto*, "below"); III) those that are directly followed by the noun phrase and do not admit any other preposition (e.g. *verso*, "towards").

behind"). [36] As accurately discussed in Meini (2010: 42 ff.), in these cases it is better to use the label "intransitive prepositions" (Klima 1965; Jackendoff 1983; Rizzi 1988: 528; Graffi 1994: 46-47; cf. also the discussion in Salvi and Vanelli 2004: 174), rather than "adverbs", or "prepositional adverbs": it is evident that the specific meaning of these lexical elements always remains a relational one (i.e. refers to the spatial relation between at least two entities).

The main reason why only this kind of preposition were included in the present study is that secondary prepositions (even the most frequent ones and those that admit the intransitive construction, such as *sotto*, *sopra*) are characterised by a semantic value[37] more restricted to the spatial domain, and from this point of view they can be considered more typical lexical words. This is a semantic property that secondary adpositions share with phrasal prepositions of the type [preposition + noun + NP] (e.g. *in mezzo a*; cf. Meini 2010: 44): for this reason, also phrasal prepositions expressing spatial relations are included in the present analysis. On the other side, primary prepositions are more polyfunctional, i.e. they generally convey a wider array of meanings (not restricted to the spatial domain) with respect to secondary prepositions.

Deictic elements (e.g. It. *lì*, *qui*, *qua*) that occur when the participant does not describe linguistically the target of the grasp, but points to (or more often touches) the pc monitor (e.g. «lo prenderei proprio qui»), were excluded from the analysis. These descriptions do not comply with the requested task (i.e. to provide a complete *verbal* desription). However, these cases are rather few.

4) **Quantity (qua)** This category collects all expressions of quantity (e.g. number words, quantifiers) such as *due*, "two", *tre*, "three, *entrambe*, "both", *tutto*, *tutti*, "all", and periphrasis expressing quantity. Such expressions often modify meronyms or spatial expressions that denote the target of the grasp (e.g. the *two* sides of the rubber boat). However, they may also be referred to the object-stimulus to be grasped (when they are related to expressions collected within the *entity* class; e.g. *un po' di farina*, "a small amount of

[36] Examples taken from Rizzi (1988: 507).

[37] «A secondary adposition (pre- or postposition) is one which expresses not a grammatical, but an objective meaning, and which may be morphologically complex and/or transparent, such as *below*, *during*. A primary adposition is one which expresses an elementary objective or a grammatical meaning and is morphologically simple, such as *of*, *in*» (Lehmann 1985: 304).

flour"); in such cases, expressions of quantity are not related to (neither constitute a reference to) the target of the grasp.[38]

5) **Perceptive (per)** All expressions (mostly adjectives) denoting perceptive properties that are referred to the target of the grasp are included in this class, such as *largo*, "large", *rosso*, "red" (e.g. *la parte larga*, "the large part", *la parte rossa*, "the red part"). Since this is not a numerous class, different kinds of perceptual properties (such as colour, shape, size, etc.) are not distinguished.

6) **Associated entity (aen)** In very few cases, the grasp described by informants is directed to an entity associated to the real target of the grasp. For instance, if the target of the grasp is a woman, the subject may choose to grasp not the woman herself (e.g. from a body part), but an accessory, for example, her bag, her scarf.

7) **Similes, based on perceptual properties (spp)** This category collects all of the words denoting a concrete entity that informants explicitly liken to the visual stimulus presented for which they have to describe a grasp. Such entities, therefore, do not constitute the target of the grasp (they were not considered in the study presented in Chapter III); however, they have been extracted for this broader analysis of grasp descriptions, since they are still referred to the object-stimulus. For example, the cup might be likened to the jug because both have handles and both are containers used to contain liquids.

The effector-related and target-related words extracted from the transcripts will be presented following the order adopted in the previous chapter (artefacts; humans; natural kinds; substances and aggregates). However, in order to allow a more in-depth qualitative analysis, each object-stimulus will be considered separately. For each stimulus, all effector-related and target-related words extracted from the descriptions will be reported in a table, together with the number of their occurrences.

[38] This argument mostly regards the grasp descriptions provided for substances and aggregates and will be fully developed in §4.6.2.

## 4.3 Artefacts

As in the previous chapter, the analysis is conducted first on the eight artefacts without affording parts, then on the artefacts with affording parts that were not presented with different orientation, and finally on the artefacts with affording parts that were presented with both a rightward and a leftward orientation.

### 4.3.1 Artefacts without affording parts

This first group of artefacts contains eight objects-stimuli: the box, the glass, the lighter, the pencil, the plate, the soccer ball, the tennis ball, and the vase.

#### 4.3.1.1 Artefacts without affording parts: detailed analysis

**BOX**

| Effector-related words | Target-related words |
|---|---|
| HAN: *mano* (1), *mani* (26).<br>SPA: *destra* (1), *sinistra* (1).<br>QUA: *due* (12), *tutte* (5), *entrambe* (9), *tutto* (1).<br>MER: *pollice* (1), *dita* (1), *palmo* (1), *palmi* (3).<br>HOL: *braccia* (7), *corpo* (2). | ENT: *oggetto* (1), *pacco* (1), *cartone* (1), *scatolone* (1).<br>SPA: *base* (2), *facce* (1), *lati* (11), *lato* (2), *sotto* (5), *lateralmente* (2), *intorno* (1), *superiore* (1), *attorno* (1), *esterni* (1), *laterale* (1), *giù* (1).<br>QUA: *due* (2).<br>PER: *corto* (1), *piccole* (1), *piccoli* (1). |

For the box, the most frequent word related to the effector is *mani*, with 26 occurrences, followed by *palmi* (only three occurrences). Clearly, frequency data suggest that the object affords a two-handed grasp. Because of its size, participants also mentioned the arms: but in most cases, the arms are indicated as only a further support for a two-handed grasp («con entrambe le mani, aiutandomi con le braccia»; «con entrambe le mani, più che altro con entrambe le braccia, metterei proprio attorno diciamo allo scatolone»). The effort required to grasp the large box is often emphasised by expressions of quantity («proprio con tutto il corpo, andrei a mettere le mani giù, entrambe le mani giù sotto il pacco»).

For what regards the target of the grasp, informants typically named either the sides of the object (*lati*) or its bottom (*sotto*, *base*).

**GLASS**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (23).<br>SPA: *destra* (4), *sinistra* (2).<br>QUA: *sola* (5), *quattro* (1), *tutta* (3), *tutte* (1), *piena* (1), *entrambe* (1).<br>MER: *pollice* (8), *dita* (8), *indice* (2), *mignolo* (1). | ENT: *bicchiere* (7), *oggetto* (1).<br>SPA: *circonferenza* (2), *lati* (1), *lato* (2), *metà* (1), *parte* (5), *sotto* (1), *dietro* (2), *davanti* (1), *frontale* (1), *opposta* (1), *bassa* (1), *attorno* (1), *intorno* (4), *retro* (1), *verso* (2).<br>QUA: *tutto* (1). |

The frequencies of effector-related words indicate that the object is mostly grasped with only one hand, either the left or the right one («una sola mano, la mano destra»). For what regards the target of the grasp, it is usually evoked by words pertaining to the spatial domain («questo lo afferro al lato»; «semplicemente, quattro dita dietro e il pollice sempre che tiene il davanti») and sometimes the object is explicitly mentioned («metterei tutta la mano intorno al bicchiere»; «con una mano, da metà bicchiere insomma»).

**LIGHTER**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (19), *pugno* (2), *polso* (1).<br>SPA: *mezzo* (1), *destra* (2).<br>QUA: *due* (3), *tre* (2), *quattro* (1), *sola* (2), *tutto* (1), *tutta* (4), *pieno* (1).<br>MER: *pollice* (11), *indice* (5), *dita* (9), *medio* (1), *palmo* (2). | ENT: *oggetto* (1), *corpo* (1), *accendino* (2).<br>SPA: *parte* (1), *intorno* (1), *sotto* (1), *dietro* (1), *davanti* (2), *sopra* (3). |

For the lighter, too, most descriptions denote a grasp performed with only one hand («con una mano sola, proprio tenendolo racchiuso dentro la mano»), but it is worth noting the high number of mentions of the fingers («con tre dita sul corpo»), probably due to the fact that the lighter affords a precision grip. Only one informant mentioned the wrist as the effector of the grasp, probably as a metonymic expression for *mano* («questo lo prenderei con tutto il polso»). Sometimes, the effector of the grasp is indicated as the object as a whole («con tutta la mano circonderei l'oggetto e stringerei»), otherwise, in few cases, informants referred to

specific parts of the object, always in spatial terms («accendino, tutta la mano, con le quattro dietro, me lo metto sul palmo, il pollice davanti»).

**PENCIL**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (9).<br>SPA: *mezzo* (1), *destra* (2), *sinistra* (1).<br>QUA: *due* (8), *tre* (3), *sola* (3).<br>MER: *pollice* (14), *indice* (12), *anulare* (1), *dita* (16), *dito* (1), *medio* (6), *punta* (2). | ENT: *matita* (1).<br>SPA: *lato* (1), *lunghezza* (1), *mezzo* (1), *parte* (3), *zona* (1), *dietro* (1), *lungo* (1), *superficie* (1).<br>MER: *punta* (1). |

As with the glass and the lighter, the pencil is also always grasped with only one hand («questa con una mano sola»), but we observe a very high frequency of mentions of the fingers (even the fingertips), because the object is very thin and affords a precision grip («con due dita, con l'indice e il pollice»; «ok, matita… sì, la prenderei con la punta delle dita, quindi le prime, indice, medio e, da dietro, il pollice»).

A few participants also specified the precise place to where the grasp would have been directed, namely, in the middle part of the object or near to its tip («con tre dita, la parte in mezzo, diciamo»).

**PLATE**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (17), *mani* (6).<br>SPA: *destra* (3).<br>QUA: *piena* (1), *sola* (2), *entrambe* (1), *ambedue* (1), *tutte* (2), *due* (5), *quattro* (2).<br>MER: *pollice* (7), *dita* (3), *indice* (4), *medio* (1), *palmo* (3). | ENT: *piatto* (2).<br>SPA: *bordo* (5), *bordi* (1), *lati* (3), *lato* (1), *estremità* (1), *fondo* (1), *parte* (2), *sotto* (8), *inferiore* (1), *superiore* (1), *sopra* (2), *lateralmente* (1).<br>QUA: *due* (1). |

Looking at effector-related words, we notice that, for the plate, participants opted partly for a one-handed grasp, partly for a two-handed grasp. In both cases, the target is the border of the plate, grasped either by one («con una mano, da un bordo») or by two sides of the dish («appoggerei tutte e due le mani sul bordo»; «con due mani, toccando i lati del piatto»).

Informants also indicated another kind of action, with the hand under the plate («questo piatto lo posso prendere con una mano, con il palmo, da sotto»): however, such descriptions do not refer properly to a grasp, but only to a way to sustain a dish.

**SOCCER BALL**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (1), *mani* (27).<br>QUA: *due* (20), *entrambe* (6), *tutte* (6).<br>MER: *palmi* (1). | ENT: *pallone* (1).<br>SPA: *lato* (2), *lati* (5), *poli* (1), *opposti* (1), *sotto* (1), *attorno* (1).<br>QUA: *tutto* (1).<br>SPP: *pallina* (1). |

Effector-related words strongly indicate that the most typical kind of grasp afforded by the soccer ball is a two-handed grasp («con tutte e due le mani»). Furthermore, in some cases, informants also referred linguistically to a grasp directed to the sides of the spherical object («con due mani, ai lati, cercando di non farlo scappare»; «palla, ai lati») or with their hands around it («il pallone lo prenderei con due mani che si chiudono attorno al pallone»), always referring to the target of the grasp in spatial terms.

**TENNIS BALL**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (29), *pugno* (1).<br>SPA: *destra* (4), *sinistra* (1).<br>QUA: *sola* (1), *intera* (2), *piena* (3), *due* (1), *tre* (1), *quattro* (1), *cinque* (1), *tutta* (10), *tutte* (3), *tutto* (1).<br>MER: *pollice* (1), *dita* (6), *palmo* (4). | ENT: *pallina* (4), *palla* (4).<br>SPA: *alto* (2), *parte* (1), *dietro* (1), *sopra* (2), *intorno* (3).<br>QUA: *tutta* (1), *completamente* (1). |

In all grasp descriptions provided by informants, only one hand is involved as the effector of the action, and in most cases participants referred to a power grasp («questa con tutta la mano, con tutte le dita»; «con una mano, l'afferrerei con una mano intera»). In one case, as already discussed for the plate, the description refers to a static posture, with the tennis ball placed on the palm of the flat hand («la potrei mettere sul palmo della mano»).

The target of the grasp is usually denoted in spatial terms, but sometimes the whole object is mentioned («con tutta la mano intorno alla palla»; «semplicemente la mano sopra la pallina e stringerei»).

**VASE**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (2), *mani* (17).<br>QUA: *entrambe* (5), *due* (13), *tre* (1), *quattro* (1), *tutte* (9).<br>MER: *dita* (3), *indice* (2), *pollice* (3), *medio* (2). | ENT: *corpo* (1), *vaso* (4), *brocca* (1), *ampolla* (1), *oggetto* (2).<br>SPA: *base* (2), *centro* (2), *circonferenza* (1), *lati* (1), *parte* (7), *sotto* (1), *inferiore* (1), *superiore* (1), *dietro* (1), *su* (1), *verso* (1), *basso* (3), *intorno* (4), *attorno* (1).<br>QUA: *due* (1).<br>MER: *collo* (8), *imboccatura* (1), *pancia* (2).<br>PER: *rotonda* (1), *rigonfiamento* (1), *stretta* (2), *sottile* (1), *larga* (1). |

Considering the words extracted from the descriptions provided by informants, the vase appears to be grasped in most cases with two hands. Only rarely did informants refer to a one-handed grasp («con una mano lo posso prendere, se non è troppo pesante»). Objects' parts are sometimes named, such as the neck («questo con entrambe le mani per il collo») or the mouth, but in most cases these parts are indicated by words pertaining to the spatial or perceptive domain («solitamente lo prenderei dal centro»; «stringendolo intorno alla parte più stretta»).

#### 4.3.1.2 Artefacts without affording parts: general considerations

We can now compare results from the analysis of effector- and target-related words for artefacts without affording parts.

| | EFFECTOR | | | | | TARGET | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | hnd | mer | spa | qua | hol | ent | mer | spa | qua | per | spp |
| Box | **27** | 6 | 2 | **27** | 9 | 4 | - | **29** | 2 | 3 | - |
| Glass | **23** | 19 | 6 | 12 | - | 8 | - | **26** | 1 | - | - |
| Lighter | 22 | **28** | 3 | 14 | - | 4 | - | **9** | - | - | - |
| Pencil | 9 | **52** | 4 | 14 | - | 1 | 1 | **10** | - | - | - |
| Plate | **23** | 18 | 3 | 14 | - | 2 | - | **27** | 1 | - | - |
| Soccer ball | 28 | 1 | - | **32** | - | 1 | - | **11** | 1 | - | 1 |
| Tennis ball | **30** | 11 | 5 | 24 | - | 8 | - | **9** | 2 | - | - |
| Vase | 19 | 10 | - | **29** | - | 9 | 11 | **27** | 1 | 6 | - |
| Tot. | **181** | 145 | 23 | 166 | 9 | 37 | 12 | **148** | 8 | 9 | 1 |
| % | **34.5** | 27.7 | 4.4 | 31.7 | 1.7 | 17.2 | 5.6 | **68.8** | 3.7 | 4.2 | 0.5 |

Table 4.1. Classification of effector-related (tot. 524) and target-related words (tot. 215) provided for the eight artefacts without affording parts.

The analysis previously conducted (§3.2.1) has shown that, for this class of objects, the number of descriptions that contain an explicit mention of the effector is high for most visual stimuli and does not show great variation. We can now make some further remarks about the semantic classification of the words extracted for the effector of the grasp.

The frequency and the classification of words referred to the effector of the grasp indicates that the reference to the hand (or the hands) is the most present in the descriptions of the vase, the tennis ball, the soccer ball, the plate, the glass, and the box; only for the pencil and the lighter does the number of mentions of the hand's meronyms exceed the number of mentions of the hand itself. This is clearly related to the object's size: the lighter and the pencil are the two smallest objects within this class of stimuli and may afford a pinch grasp. On the contrary, meronymic expressions are rarer for the soccer ball and the box, which are large objects; moreover, only for the box (which is the largest object-stimulus) are the arms also named as the effectors of the grasp, together with the two hands.

Expressions of quantity are especially frequent for the box, the soccer ball, and the vase, and in most cases are due to the fact that two effectors (both the right and the left hands) are involved in the action (cf. the high frequency of words such as *due*, *entrambe* for these objects).

For what regards the target of the grasp, we have already observed (§3.2.1) that the number of references to the target varies a lot among the different objects-stimuli: it is particularly

low for the pencil, the lighter, the tennis ball, and the soccer ball, but it is higher for the box and the vase.

For all objects-stimuli, most words referred to the target of the grasp pertain to the spatial domain. However, such kinds of lexical expressions are less frequently produced for the spherical objects (the tennis ball and the soccer ball) and the smallest ones (the lighter and the pencil), for which internal subspaces, such as the sides, the upper or the lowest part, etc. are less salient, or more difficult to individuate (especially for the two balls, which have a continuous surface).

The main characteristic of this first class of objects-stimuli is that they do not have any specific part designed to facilitate grasping (such as a handle). Nevertheless, objects' parts are sometimes named for the pencil and especially for the vase, which has a more complex shape (cf. §3.2.1) and is constituted by identifiable parts, some of which are particularly suitable for grasping (such as the neck) and can be named by informants.

### 4.3.2 Artefacts with affording parts (without different orientation)

In this category, we find artefacts that either have one part typically involved in the grasp and designed for it (the coffee cup, the backpack, the girl bag, the rubber boat, the trolley, the umbrella), or have more than one prominent part suited for grasping (the chair).

#### 4.3.2.1 Artefacts with affording parts (without different orientation): detailed analysis

**CHAIR**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (5), *mani* (18).<br>SPA: *destra* (1), *sinistra* (1).<br>QUA: *due* (8), *otto* (1), *tutte* (2), *entrambe* (3).<br>MER: *pollice* (1), *pollici* (1), *dita* (4), *palmi* (1).<br>HOL: *braccia* (1). | ENT: *sedia* (3).<br>SPA: *alta* (1), *base* (1), *basso* (2), *esterne* (1), *estremità* (1), *lati* (2), *metà* (1), *parte* (11), *sotto* (7), *inferiore* (1), *superiore* (3), *dietro* (4), *davanti* (1), *laterali* (1), *sopra* (3), *verticali* (1), *lateralmente* (1).<br>MER: *aste* (1), *buchi* (1), *spazio* (1), *cuscino* (2), *gambe* (1), *schienale* (12), *sedile* (1), *seduta* (1), *spalliera* (3), *spalline* (1), *stecche* (1).<br>QUA: *due* (2). |

The chair is generally grasped with two hands. Most informants said that they would have directed their grasp to a specific part of the object: frequently, the back of the chair; otherwise, its seat or its legs («potrei prenderla dallo schienale, con le mani ai lati dello schienale»). We observe that the vast majority of grasps are directed to the upper part of the object: this is the part most suited for the grasping because it usually reaches human hands. The middle part, i.e. the seat, is chosen more rarely, whereas the lowest one (the support), which is reachable only with difficulty, is mentioned only once («dalle gambe sotto, con le mani»). The holes in the backrest were considered as parts of the chair: some informants explicitly named them as the parts more suited for the grasping («dalla spalliera, mettendo le dita tra gli spazi verticali»).

In one case, the seat is indicated as the "real chair", in contrast with the back («mettendo quindi la mano destra sullo schienale, la mano sinistra sotto la sedia vera e propria»).

**COFFEE CUP (OVERTURNED)**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (19), *mani* (1).<br>SPA: *destra* (2), *sinistra* (1), *interno* (1), *verso* (2), *basso* (2).<br>QUA: *due* (2), *cinque* (1), *tutte* (5), *tutta* (1).<br>MER: *pollice* (1), *indice* (1), *punta* (1), *dita* (13), *palmo* (2).<br>SPP: *gru* (1), *conca* (1). | ENT: *corpo* (1), *tazzina* (1), *tazza* (4).<br>SPA: *alto* (4), *base* (1), *cerchio* (1), *finale* (1), *fondo* (4), *lati* (3), *lato* (1), *parte* (4), *dentro* (1), *sotto* (1), *intorno* (3), *minore* (1), *inferiore* (3), *sopra* (6), *basso* (2), *cima* (1).<br>MER: *manico* (2).<br>PER: *stretta* (1).<br>QUA: *due* (1).<br>SPP: *pallina* (1). |

For the overturned coffee cup, we observe a strong tendency to describe a grasp performed with the whole hand («questa la prenderei con tutta la la mano e le dita intorno alla tazza»), most frequently directed to the bottom of the cup, i.e. the upper part of the overturned object («allora questo lo prenderei con una mano da sopra, dalla parte più stretta in cima»; «non so, da sotto, cioè sì da sopra, in questo caso dal fondo»). The handle is named as the target of the grasp in very few cases («sempre con pollice e indice intorno al manico»). It is worth noting that in two cases the shape of the hand is described by referring to a cup and a crane («con la mano messa a conca, quindi… però con il palmo rivolto verso il basso,

e la prenderei per il fondo e la solleverei»), and in one case the coffee cup is likened to a tennis ball.

Although this object-stimulus presents an affording part, i.e. a part specifically designed for grasping, informants largely ignore it. It seems that, when the cup is overturned, the handle loses in part its capacity to attract the grasp (we could say its "affording power"). In fact, we usually take cups by their handle when they are filled with some liquid and we are going to drink their content. However, when cups are overturned, we cannot use them straight for drinking, whereas, for instance, we are likely to put them in the cupboard to keep them clean, or to leave them on the draining board to dry off. Besides the influence of possible subsequent actions on the kind of grasp described by informants (even in front of a static picture and without any request to act), we should consider also the small dimension of the object. For a large cup, the handle would offer a more economic, comfortable, and firm grasp, with respect either to a grasp with a single hand stretched to hold the object, or to a two-handed grasp. However, in this case, taking the handle of the small overturned coffee cup would require a very precise and controlled hand shaping, which is far from leading to any benefit or advantage, except when there is the need for drinking from it (and in this case the cup would be overturned again).

**GIRL BAG**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (11), *pugno* (1). | ENT: *borsa* (2). |
| QUA: *sola* (1), *quattro* (1). | SPA: *sotto* (1), *intorno* (2). |
| MER: *pollice* (1), *dita* (3), | MER: *bretella* (1), *manica* (1), *manico* (23), *maniglie* (1), *manici* (6), *tracolla* (1). |
| *falangi* (1). | QUA: *due* (1), *tutti* (1). |
| HOL: *braccio* (1). | SPP: *zaino* (1). |

For the girl bag, all of the participants described a grasp directed to the handle, mostly performed with one hand. There is no other kind of grasp referred. The only difference that emerged regards the number of handles mentioned by the informants, either one or two («la borsa la prenderei dal manico»; «prendendo entrambi i manici, con una sola mano»). Probably due to the presence of a top handle, the bag is likened to the backpack («dal manico, come lo zaino»).

In many cases, the effector of the grasp is not named; otherwise, participants indicate a single hand or the fingers («qui la borsa intuitivamente stringendo il pugno dal manico»; «la borsa l'afferrerei per il manico, per la tracolla, con le dita intorno»).

**RUBBER BOAT**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (9), *mani* (6).<br>SPA: *destra* (1).<br>QUA: *due* (3), *entrambe* (2), *tutte* (3), *intero* (1).<br>MER: *dito* (1), *dita* (1).<br>HOL: *braccio* (3), *braccia* (2). | ENT: *canotto* (2).<br>SPA: *lati* (1), *lato* (1), *bordi* (1), *bordo* (1), *destra* (1), *sinistra* (1), *parte* (3), *punta* (1), *interno* (2), *dentro* (1), *sotto* (2), *intorno* (2), *superiore* (1).<br>MER: *manici* (5), *maniglie* (2), *maniglia* (1), *corde* (3), *corda* (2), *cordino* (1), *cordoncini* (1), *filo* (2), *elastici* (1), *passantine* (1).<br>PER: *lunghi* (1), *gonfia* (1).<br>QUA: *due* (1). |

For the rubber boat, in most cases participants described a grasp directed to a part of the object, namely, the handles, i.e. the part more suited for the grasping («dagli appositi manici, con tutte e due le mani»), but also the rope that surrounds it («afferrando il canotto con le corde che ha per tirarlo»), or the rings where the rope is inserted. Otherwise, the inflated sides are a good target, grasped with two hands and sometimes also with the arms («lo prenderei dai lati, quelli più più lunghi, con entrambe le mani»; «con due mani, però se è troppo largo probabilmente mi serviranno anche le braccia»).

**BACKPACK**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (10), *mani* (3).<br>QUA: *due* (2), *entrambe* (1), *tutte* (1), *sola* (2), *solo* (1).<br>MER: *palmo* (1), *dita* (1).<br>HOL: *braccio* (1). | ENT: *zaino* (2), *oggetto* (1).<br>SPA: *alta* (1), *alto* (2), *finale* (1), *lati* (1), *parte* (6), *dietro* (1), *sopra* (5), *superiore* (5).<br>MER: *aggancio* (1), *attacco* (1), *attaccatura* (1), *braccioli* (1), *bretella* (1), *bretelle* (1), *cinghia* (1), *cinghie* (1), *fascetta* (1), *fibbia* (1), *gancetto* (1), *lacci* (1), *laccio* (5), *manico* (5), *maniglia* (2), *tracolle* (1).<br>PER: *azzurro* (1), *piccolo* (1).<br>QUA: *due* (1).<br>SPP: *valigetta* (1). |

Most of the thirty descriptions provided for the backpack are performed with only one hand and directed to a specific part of the object, either the top handle («dal manico superiore, stringendo le dita») or the shoulder straps («dalle bretelle, usando tutte e due le mani»).

Words denoting the object's parts are the most frequent; however, expressions pertaining to the spatial domain are also found: for instance, two informants mentioned that they would have grasped the object with two hands, from its sides («lo prenderei con entrambe le mani dai lati»). It is interesting to note that, due to the presence of the handle, the backpack is likened to a briefcase («come prima cosa, lo afferrerei dal laccio che è in alto, come se dovessi tirar su una valigetta»).

**TROLLEY**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (6), *mani* (4), *pugno* (1).<br>SPA: *destro* (1).<br>QUA: *due* (2), *entrambe* (2), *quattro* (1), *tutta* (1), *sola* (1).<br>MER: *dita* (2).<br>HOL: *braccia* (1). | ENT: *oggetto* (1).<br>SPA: *parte* (3), *sotto* (1), *superiore* (7), *dietro* (1), *alta* (1), *cima* (1), *bordo* (1), *lati* (1), *davanti* (1), *sopra* (3), *lateralmente* (1), *attraverso* (1), *intorno* (1).<br>MER: *cinghietta* (1), *fascetta* (1), *manico* (20), *pezzettino* (1), *maniglia* (2). |

Most of the thirty descriptions collected for the trolley contain a mention of the handle, which is the word most frequently named («la valigia la prenderei per il manico, con le dita intorno al manico stesso»; «sempre dal manico»).

On the contrary, the effector of the grasp is rarely mentioned; only in six cases do we find *mano* («dal manico superiore, usando tutta la mano»), whereas the mentions of the hand's parts are even rarer. Only a few informants described a two-handed grasp («andrei con due mani sul bordo dell'oggetto»).

**UMBRELLA**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (15), *mani* (2), *pugno* (1).<br>SPA: *destra* (3), *sinistra* (1).<br>QUA: *due* (1), *entrambe* (1),<br>*quattro* (1), *tutta* (2), *sola* (1).<br>MER: *pollice* (1), *dita* (3), *palmo* (1). | ENT: *ombrello* (5).<br>SPA: *alto* (1), *estremità* (1), *finale* (1), *lato* (1), *metà* (1), *mezzo* (1),<br>*basso* (1), *intorno* (4), *parte* (4).<br>QUA: *due* (1).<br>MER: *manico* (18), *impugnatura* (4), *tela* (1), *fusto* (1), *corpo* (1), *collino* (1).<br>PER: *verde* (1), *ricurvo* (1). |

For the umbrella, as for the trolley, most participants described a one-handed grasp directed to the handle («dall'impugnatura, con tutta la mano»; «questo, allora, con una mano dal manico»). Again, *manico* ("handle") is the most frequent target-related word. In five answers, participants mentioned only the handle («per il manico»; «dal manico»).

Otherwise, the object can be grasped by its stem («con una mano, da metà ombrello»; «posso prenderlo dal mezzo, col palmo, stringendo con le dita»).

It is interesting to observe that in one case the umbrella is identified with its long narrow part, in contrast with the handle («allora questo di solito lo prendo non dal manico, ma proprio dall'ombrello, con una mano»).

#### 4.3.2.2 Artefacts with affording parts (without different orientation): general considerations

Table 4.2. summarises results from the analysis of effector- and target-related words for artefacts with affording parts, not presented with different orientation.

In the previous analysis (§3.2.2), we observed that, for this group of artefacts with affording parts, there are, in general, more references to the target of the grasp with respect to the number of references to the effector (except for the overturned coffee cup). Looking in more detail at the types of words related to the effector, we notice that words referring to the hand are the most frequent for all objects; however, the number of mentions of the fingers is particularly high for the overturned coffee cup, which is the smallest object-stimulus within this category.

| | EFFECTOR | | | | | | TARGET | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | hnd | mer | spa | qua | hol | spp | spa | mer | qua | ent | per | spp |
| Chair | **23** | 7 | 2 | 14 | 1 | - | **42** | 25 | 2 | 3 | - | - |
| Coffee cup | **20** | 18 | 8 | 9 | - | 2 | **37** | 2 | 1 | 6 | 1 | 1 |
| Girl bag | **12** | 5 | - | 2 | 1 | - | 3 | **33** | 2 | 2 | - | 1 |
| Rubber boat | **15** | 2 | 1 | 9 | 5 | - | 18 | **19** | 1 | 2 | 2 | - |
| Backpack | **13** | 2 | - | 7 | 1 | - | 22 | **25** | 1 | 3 | 2 | 1 |
| Trolley | **11** | 2 | 1 | 7 | 1 | - | 23 | **25** | - | 1 | - | - |
| Umbrella | **18** | 5 | 4 | 6 | - | - | 15 | **26** | 1 | 5 | 2 | - |
| Tot. | **112** | 41 | 16 | 54 | 9 | 2 | **160** | 155 | 8 | 22 | 7 | 3 |
| % | **47.9** | 17.5 | 6.9 | 23.1 | 3.8 | 0.8 | **45.1** | 43.6 | 2.3 | 6.2 | 2 | 0.8 |

Table 4.2. Classification of effector-related (tot. 234) and target-related words (tot. 355) provided for the seven artefacts with affording parts, presented without different orientation.

Considering the words extracted with relation to the target, some differences among objects emerge. In particular, meronymic expressions (that mostly consist of mentions of an object's handle) are generally more frequent than words pertaining to the spatial domain. However, this is particularly evident for the girl bag, for which we register the highest number of words denoting the object's meronyms and the lowest number of words related to to the spatial domain.

As the analysis conducted for each object has revealed, the number of mentions of an object's parts reflects a strong and clear tendency to prefer the handle, i.e. the affording part, as the target of the action. Words denoting the handle (the lemmas *manico*, *impugnatura*, *maniglia*) are 92, i.e. 59% over the total number of words related to meronyms. When two or more different object's parts compete (as for the backpack and, marginally, for the rubber boat), other meronymic expressions occur, but the affording part is still preferred. The only stimulus that is not aligned to this pattern is the overturned coffee cup, for which in most cases participants named a power grasp from the above (or laterally), similar to the grasp of a tennis ball, and did not mention the handle.

### 4.3.3 Artefacts with affording parts (with different orientation)

For the eight artefacts with affording parts that during the experiment were presented with different orientation (the hairdryer, the jug, the ladle, the microphone, the pacifier, the

sword, and the tea-cup), each stimulus will be analysed separately; but, as already done in Chapter III, the content of the descriptions provided by the two groups of informants (R-H, i.e. right-handed vs. L-H, i.e.left-handed) will be also compared.

#### 4.3.3.1 Artefacts with affording parts (with different orientation): detailed analysis

**HAIRDRYER (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (12).<br>SPA: *destra* (1).<br>QUA: *tutta* (1), *tutto* (1), *sola* (2), *quattro* (1), *tutte* (2), *intera* (1).<br>MER: *pollice* (2), *dita* (4), *palmo* (1). | ENT: *phon* (4), *oggetto* (1).<br>SPA: *lato* (1), *intorno* (2), *parte* (6), *inferiore* (1).<br>MER: *impugnatura* (6), *manico* (12).<br>PER: *lunga* (1), *stretta* (1). |
| L-H | HND: *mano* (3).<br>SPA: *sinistra* (1).<br>QUA: *tutta* (1).<br>MER: *dita* (2), *palmo* (1). | SPA: *lato* (1), *intorno* (1).<br>MER: *manico* (7). |

For what regards effector-related words, we observe that in both groups of informants the reference is always to only one hand, sometimes indicated as a "whole" hand («il phon, mano intera, stringerei il manico del phon»), whereas the hand's meronyms are rarely indicated («semplicemente lo impugnerei e lo tirerei su, quindi normalmente, con tutte le dita»).

The most frequent target-related words are *manico* and *impugnatura*, both meaning "handle" («dall'impugnatura, con tutta la mano»). Sometimes, the reference to the object (and in most cases to the handle) is made through words pertaining to the perceptive or spatial domain («impugnerei la parte più lunga e più stretta, di lato»).

We should also remark that in one case a left-handed informant described a grasp directed to the rightward-oriented handle, but performed with his dominant (left) hand («lo prendo per il manico, con la mano sinistra tendenzialmente, avvolgendo le dita intorno al manico»; «con una mano, dalla parte inferiore»).

**HAIRDRYER (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (18).<br>SPA: *destra* (2), *sinistra* (4).<br>QUA: *sola* (1), *intera* (1), *quattro* (1), *tutta* (1).<br>MER: *dita* (2), *pollice* (2), *palmo* (1). | ENT: *phon* (1).<br>SPA: *parte* (3), *sotto* (1), *sopra* (1), *lato* (2), *intorno* (3).<br>MER: *impugnatura* (2), *manico* (13).<br>PER: *concava* (1), *stretta* (1). |
| L-H | HND: *mano* (2).<br>SPA: *sinistra* (1).<br>QUA: *solo* (1).<br>MER: *dita* (1), *palmo* (1). | SPA: *intorno* (1).<br>MER: *manico* (7). |

In the descriptions provided for the leftward-oriented hairdryer, the object is grasped in most cases with only one hand (there are no occurrences of the plural word *mani*), by its handle («il phon, con una sola mano intorno al manico»). Sometimes, the target of the grasp is indicated with adjectives pertaining to the spatial domain, or to the domain of visual perception («con una mano, di lato, nella parte più stretta»; «sicuramente metterei il pollice sotto, nella parte concava diciamo, e la mano sopra»).

For what regards the effector of the grasp, we observe that in five cases (predominantly within the right-handed group) informants specified that they would have used the left hand to grasp the object («questo lo prenderei con la mano sinistra dal manico»).

**JUG (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (15), *mani* (2), *pugno* (1).<br>SPA: *destra* (6) *sinistra* (1).<br>QUA: *due* (2), *quattro* (3), *cinque* (1), *tutte* (1).<br>MER: *dita* (4), *pollice* (3). | ENT: *brocca* (3).<br>SPA: *parte* (4), *sotto* (1), *dietro* (1), *sopra* (1), *intorno* (2), *destra* (1), *posteriore* (1), *base* (1).<br>MER: *impugnatura* (1), *manica* (1), *manico* (16), *occhiello* (1). |
| L-H | HND: *mano* (6).<br>SPA: *destra* (2).<br>QUA: *tutta* (1).<br>MER: *dita* (2). | SPA: *parte* (1), *superiore* (1), *intorno* (1).<br>MER: *manico* (6). |

For the rightward-oriented jug, most descriptions contain an explicit reference to a single hand as the effector of the grasp, and to the handle as the target («con la mano destra,

afferrerei il manico»). However, spatial terms are sometimes used with reference to the jug, and in most cases they are referred to the handle («con la mano dalla parte destra della caraffa», where "the right part of the jug" clearly refers to the handle).

Only in few cases did participants describe a two-handed grasp («con la mano destra prenderei il manico, con la sinistra la terrei sotto»; «con tutte e due le mani, adagiandole sulla base»).

It is worth noting that eight informants (two are in the left-handed group) specified that they would have used the right hand to perform the grasp («la prenderei per il manico, con la mano destra, nonostante io sia mancino»).

**JUG (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (10), *mani* (4), *pugno* (1).<br>SPA: *destra* (2), *sinistra* (8).<br>QUA: *tutta* (1), *tutte* (1), *due* (5), *quattro* (1).<br>MER: *pollice* (2), *dita* (3). | ENT: *caraffa* (1).<br>SPA: *lato* (2), *parte* (1), *sotto* (1), *sopra* (1), *intorno* (3).<br>MER: *collo* (4), *impugnatura* (2), *manico* (11). |
| L-H | HND: *mano* (3).<br>SPA: *sinistra* (2).<br>MER: *dita* (1). | SPA: *davanti* (1), *intorno* (1).<br>MER: *manico* (5). |

Similarly, the leftward-oriented jug is also predominantly grasped with only one hand, by the handle («la prenderei dal manico, usando tutta la mano»). Only in a few cases did informants describe a two-handed grasp that sometimes still involved the handle («metterei una mano di sotto e l'altra la metterei al manico»); otherwise, it is directed to the neck of the jug («con tutte e due le mani dal collo»).

Since this jug is leftward-oriented, ten participants (eight of which are right-handed) mentioned that they would have used the left hand («con la mano sinistra, perché se è girato di qua, con la sinistra, cioè perché mi viene spontaneo farlo così»). Only two right-handed informants explicitly named the right hand as the effector of the grasp of the leftward-oriented jug («la prenderei comunque con la destra»).

**LADLE (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (11).<br>SPA: *destra* (8).<br>QUA: *quattro* (1), *sola* (2).<br>MER: *pollice* (4), *indice* (3), *medio* (2), *dita* (4). | ENT: *mestolo* (1).<br>SPA: *cima* (2), *destra* (1), *destro* (1), *estremità* (1), *lato* (1), *metà* (1), *parte* (3), *sotto* (2), *intorno* (1).<br>MER: *manico* (19), *impugnatura* (2), *cucchiaio* (1).<br>SPP: *penna* (1). |
| L-H | HND: *mano* (4).<br>SPA: *destra* (1), *sinistra* (2).<br>QUA: *solo* (1).<br>MER: *pollice* (2), *indice* (2), *medio* (2), *dita* (2). | ENT: *mestolo* (1).<br>SPA: *parte* (1), *alta* (1).<br>MER: *manico* (5). |

Looking at effector-related words, first of all, we notice that in the descriptions provided for the rightward-oriented ladle, the object always appears to be grasped with only one hand («questo sempre con solo una mano, dal manico»). However, with comparison to other artefacts with affording parts, fingers are frequently mentioned, probably because of the thin shape of the handle that, in one case, is likened to a pen («le dita indice, medio e pollice»; «con la mano destra, come una penna, appoggiando sul medio, e indice e pollice che lo fermano»).

Considering target-related words, it is clear that the handle of the ladle is the preferred target and *manico*, *impugnatura* are the words most frequently named («con una mano, dalla parte del manico»; «per il manico, sì, sempre con la mano destra»), even if among meronyms we find also *cucchiaio*, "spoon", denoting the opposite part with respect to the handle. Words pertaining to the spatial domain are often used in association with meronyms, to specify which part of the long handle of the ladle is the target of the grasp («da metà manico»; «mettendo il manico dalla parte in cima fra pollice, indice e medio»).

Most descriptions containing an indication of which hand would be involved in the action of grasping refer to the right hand; however, two left-handed informants preferred their dominant hand («in questa posizione esatta, probabilmente lo prenderei con la mano sinistra»).

**LADLE (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (18), *pugno* (1).<br>SPA: *destra* (5), *sinistra* (1).<br>QUA: *tutte* (2), *tutta* (2).<br>MER: *pollice* (3), *indice* (3), *medio* (1), *dita* (4). | ENT: *mestolo* (4).<br>SPA: *estrema* (1), *iniziale* (1), *verso* (1), *vicino* (1), *sinistra* (2), *metà* (1), *parte* (7), *fondo* (1), *alta* (1), *sotto* (1), *superiore* (1), *intorno* (3).<br>MER: *manico* (17), *impugnatura* (2).<br>SPP: *penna* (2). |
| L-H | HND: *mano* (2).<br>SPA: *sinistra* (1).<br>QUA: *due* (1)<br>MER: *pollice* (2), *indice* (2), *medio* (1), *dita* (3). | SPA: *esterna* (1), *parte* (2), *alta* (1).<br>MER: *asta* (1), *manico* (4).<br>PER: *lunga* (1). |

For the leftward-oriented ladle, we may observe (when information about the effector is provided) that participants always described a one-handed grasp that in most cases is explicitly directed to the handle («prenderei per il manico, anche questo premendo le dita intorno al manico»). Also, in this orientation condition, fingers are frequently mentioned and the ladle in two cases is likened to a pen[39] («metterei la mano sull'impugnatura, chiudendo tra indice e pollice»; «prenderei il manico, anche lì stringendo come stringerei una penna più o meno, cioè farei passare il manico fra l'indice il medio e il pollice»; «questo mestolo lo prenderei con la mano sinistra, in particolare con le dita, penso pollice indice e medio»).

Five informants specified that they would have used their right hand in order to grasp the leftward-oriented ladle. Notably, they are all right-handed participants that preferred to choose as effector their dominant hand, rather than the hand that was spatially aligned with the target of the grasp («lo prenderei sempre con la mano destra, però farei un movimento rotatorio, dal momento che c'è il manico del mestolo a sinistra»).

[39] In one case, it was the same informant that stated the same comparison both for the rightward- and the leftward-oriented object.

**MICROPHONE (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (12).<br>SPA: *destro* (1), *destra* (2).<br>QUA: *due* (2), *quattro* (1), *tutta* (1), *tutte* (1), *sola* (3).<br>MER: *pollice* (1), *medio* (1), *dita* (4), *anulare* (1), *mignolo* (1). | ENT: *microfono* (2), *oggetto* (1).<br>SPA: *verso* (1), *finale* (1), *fondo* (1), *lato* (1), *cima* (1), *metà* (1), *parte* (8), *sotto* (1), *lateralmente* (1), *basso* (1), *intorno* (1).<br>MER: *manico* (4), *impugnatura* (4), *gambo* (1).<br>PER: *spessa* (1), *fine* (2), *stretta* (1). |
| L-H | HND: *mano* (2), *pugno* (1).<br>QUA: *solo* (1).<br>MER: *palmo* (1), *dita* (1). | SPA: *centrale* (1), *parte* (2), *intorno* (1).<br>MER: *manico* (4), *impugnatura* (1), *tasto* (1). |

Considering the effector-related words extracted by descriptions provided by informants, the rightward-oriented microphone is mostly grasped with a single hand («microfono, con una sola mano»). The fingers are rarely mentioned, but always with reference to a power grasp («chiudendo tutte le dita intorno al microfono»). Only in 14 cases did informants explicitly mention the handle of the microphone, whereas visuo-spatial expressions are generally preferred («il microfono, stringo il manico del microfono con una mano»; «con la mano destra, lo stringerei e basta, sì, per la parte più fine»; «sulla parte meno spessa»; «qui per la parte centrale, dove ci può essere il tasto»).

Only three right-handed informants specified that they would have used the right hand to perform the grasp.

**MICROPHONE (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (15), *pugno* (1).<br>SPA: *destra* (2), *sinistra* (7), *opposto* (1).<br>QUA: *tutta* (2), *tutte* (1), *piena* (1), *sola* (3).<br>MER: *pollice* (1), *dita* (3). | ENT: *microfono* (2), *oggetto* (2), *corpo* (1).<br>SPA: *lato* (1), *parte* (5), *sotto* (1), *intorno* (2).<br>MER: *manico* (6), *impugnatura* (4), *gambo* (1).<br>PER: *stretta* (1).<br>SPP: *phon* (1). |
| L-H | HND: *mano* (3).<br>SPA: *sinistra* (3).<br>MER: *palmo* (1), *dita* (1). | SPA: *sinistra* (1).<br>MER: *manico* (4). |

As already noticed for the rightward orientation condition, for the leftward-oriented microphone, too, there are only 15 references to the handle of the object («microfono, stringo il pugno sul manico del microfono»; «con tutta la mano sull'impugnatura»). In one case, the microphone is likened to the hairdryer.

The grasps described are always performed with only one hand, but ten informants answered that they would have preferred to use the left hand («con la sinistra dal manico»; «con la mano sinistra, dal manico diciamo»). Notably, seven of these participants are right-handed. For them, the choice of the left hand is only due to an effect of spatial alignment between the orientation of the handle and the side of the effector.

**PACIFIER (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (9).<br>SPA: *destra* (3).<br>QUA: *solo* (2), *due* (5), *tre* (1), *tutta* (1), *tutte* (1), *sola* (3).<br>MER: *pollice* (2), *indice* (3), *dito* (3), *dita* (9). | ENT: *ciuccio* (1).<br>SPA: *finale* (1), *interno* (1), *parte* (11), *sotto* (1), *superiore* (1).<br>MER: *anello* (1), *gancetto* (1), *laccio* (1), *impugnatura* (2), *manico* (7), *occhiello* (1).<br>PER: *tonda* (1), *rossa* (1). |
| L-H | HND: *mano* (2).<br>SPA: *destra* (1).<br>QUA: *due* (1).<br>MER: *pollice* (1), *indice* (1), *dita* (2). | SPA: *esterno* (1), *bordo* (1).<br>MER: *manico* (3), *occhiellino* (1). |

For the rightward-oriented pacifier, most effector-related words are referred to a single hand («con tutta la mano, sì»), but also to fingers, probably because the object's size affords a precision grip («questo, anche qui, lo prenderei con due dita»).

Among target-related words, we observe a high number of different words all referred to the little handle of the pacifier, among which the most frequent one is *manico* («con la mano destra, sempre dal manico»; «questo sempre lo prenderei dall'occhiellino, per motivi di igiene»).

Words pertaining to the domain of space or expressing perceptive properties are rather rare, and in most cases refer to the handle («il ciuccio lo prenderei dal di sotto»; «il ciuccio, dalla parte finale»; «la parte quella tonda, rossa, mettendola fra pollice e indice»).

**PACIFIER (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (8).<br>OBP: *bocca* (1).<br>QUA: *due* (4), *tre* (1), *piena* (1), *solo* (1), *tutta* (3), *sola* (1).<br>MER: *pollice* (2), *indice* (2), *medio* (1), *dito* (2), *dita* (6). | ENT: *ciuccio* (2).<br>SPA: *opposta* (1), *finale* (1), *parte* (7), *sotto* (2), *intorno* (1).<br>MER: *aggancio* (1), *anellino* (1), *gancino* (1), *mammella* (1), *manichino* (1), *manico* (6), *tondino* (1).<br>PER: *tonda* (1). |
| L-H | QUA: *due* (1).<br>MER: *pollice* (1), *indice* (1), *dita* (2). | SPA: *dietro* (1).<br>MER: *manichino* (2), *manico* (4), *occhiello* (1). |

For the leftward-oriented condition, this object appears to be grasped with only one hand; in particular, with the fingers («questo lo prenderei con tutta la mano, piena, intorno al ciuccio»; «allora, il ciuccio lo prendi con due dita dal manico»). Again, we observe many different expressions, apart from *manico*, used to indicate the handle, which constitutes the preferred target of the grasp («ciuccio, allora, lo prendo dalla parte dove c'è il tondino»; «prendendo, non so, fra pollice e indice la parte tonda»; «questo lo prenderei qua, da questo occhiello»).

**SWORD (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (11), *mani* (3), *pugno* (1).<br>SPA: *destra* (1).<br>QUA: *due* (3), *quattro* (1), *tutta* (2), *intero* (1), *tutte* (2).<br>MER: *pollice* (3), *dita* (4), *palmo* (2). | ENT: *pugnale* (1).<br>SPA: *fondo* (1), *intorno* (3), *parte* (5), *lateralmente* (1).<br>MER: *impugnatura* (9), *elsa* (2), *punta* (1), *manico* (11). |
| L-H | HND: *mano* (4).<br>QUA: *tutta* (1).<br>MER: *dita* (1), *palmo* (1). | SPA: *intorno* (1).<br>MER: *impugnatura* (2), *elsa* (1), *manico* (4). |

For the rightward-oriented sword, we observe mentions of both the hand and the fingers, but always with reference to a power grasp («dall'impugnatura, con tutta la mano»; «questa prendendo l'impugnatura, stringendola fra il pollice e le altre dita»). Since the object seems to be heavy, sometimes the use of both hands is required («credo che sia pesante, sembra,

quindi con tutte e due le mani dal manico»). The number of words related to the handle of the sword is particularly high, because the object does not afford any other kind of grasp, due to reasons of safety («questa sicuramente con una mano dall'impugnatura, senza ombra di dubbio»).

**SWORD (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (11), *mani* (1), *pugno* (2).<br>SPA: *sinistra* (7).<br>QUA: *due* (1), *tutta* (1), *tutte* (1).<br>MER: *dita* (2), *palmo* (1). | ENT: *spada* (2).<br>SPA: *parte* (2), *sopra* (2), *lateralmente* (1), *intorno* (1), *attorno* (1).<br>MER: *impugnatura* (7), *elsa* (1), *lama* (1), *manico* (9). |
| L-H | HND: *mano* (2).<br>SPA: *destra* (1), *sinistra* (1).<br>MER: *dita* (1), *palmo* (1). | MER: *impugnatura* (2), *elsa* (1), *lama* (1), *manico* (3). |

For the leftward-oriented sword, too, we observe a very high number of references to the handle of the object («con l'impugnatura»; «l'afferrerei con una mano dal manico»), whereas spatial expressions used to indicate the target are rather rare («con una mano, lateralmente»). Surprisingly, some participants also mentioned the blade; in one case, this answer was given by a right-handed informant, and was therefore the result of a spatial alignment effect («istintivamente mi verrebbe di prenderla dalla lama»); in another case, however, it came from a left-handed subject («il primo istinto sarebbe di afferrarlo per la lama, sollevarlo con la destra»).[40] However, most spatial expressions related to the effector suggest that informants (especially the right-handed ones) described a grasp performed with the left hand («afferrerei la spada dalla parte del manico con la sinistra»).

[40] It should be noted that the same subject decleared himself to be left-handed; for instance, for the leftward-oriented microphone, he answered: «sarei tentato di afferrarlo subito con la sinistra, perché sono mancino».

**TEA-CUP (RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (6), *mani* (3).<br>SPA: *destra* (2).<br>QUA: *due* (10), *tre* (2), *tutte* (2), *solo* (2), *solamente* (1).<br>MER: *dita* (10), *dito* (4), *pollice* (6), *indice* (8), *medio* (1). | ENT: *tazza* (4), *tazzina* (1).<br>SPA: *parte* (2), *dentro* (2), *destra* (1), *sotto* (1), *sopra* (1), *dietro* (1), *intorno* (3).<br>MER: *manico* (14), *manica* (1). |
| L-H | HND: *mano* (2).<br>SPA: *destra* (2).<br>QUA: *solo* (1).<br>MER: *dito* (2), *pollice* (3), *indice* (4), *medio* (1). | ENT: *corpo* (1).<br>SPA: *interno* (1), *parte* (1), *sopra* (1), *intorno* (1).<br>MER: *manico* (5), *occhiello* (1). |

The tea-cup is the only object-stimulus presented with four different orientations. In this first case, the object is upward and rightward-oriented. Considering effector-related words, we observe that spatial expressions referred to the effector, in the descriptions provided by both groups of informants, always indicate the right hand («questo sempre per il manico con la destra»). In many cases, participants mentioned the fingers: this is because the target of the grasp, in most descriptions, is the handle of the cup, which may afford a precision grip («questa tazzina di nuovo la prenderei con la mano destra, con il mio pollice e indice»; «con le tre dita della mano prenderei il manico»).

The most frequent target-related word is *manico*, often combined with spatial expression («l'afferrerei con tutte e due le mani, con un dito dentro l'apposito manico»), but there are also two other types of grasps described. A few informants described a simple two-handed grasp («questo solitamente lo prendo con due mani»).

**TEA-CUP (LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (4), *mani* (6).<br>SPA: *destro* (1), *destra* (1), *sinistra* (5).<br>QUA: *due* (9), *tutte* (4), *entrambe* (1).<br>MER: *dita* (8), *dito* (1), *pollice* (2), *indice* (3), *medio* (1). | ENT: *tazza* (1).<br>SPA: *alto* (1), *verso* (1), *sinistra* (2), *lato* (4), *opposto* (1), *intorno* (1), *parte* (3), *dentro* (1), *sotto* (1), *sopra* (1), *laterale* (1).<br>MER: *manico* (12), *impugnatura* (1), *manica* (1), *cerchiettino* (1).<br>PER: *grossa* (1).<br>SPP: *brocca* (2). |
| L-H | HND: *mano* (3).<br>SPA: *destra* (1), *sinistra* (2).<br>QUA: *due* (2).<br>MER: *dita* (2), *dito* (1), *pollice* (1), *indice* (1), *medio* (1). | SPA: *sinistra* (1), *base* (1), *lato* (1).<br>MER: *manico* (6), *occhiello* (1). |

For the same tea-cup, with the leftward-oriented handle, we observe again that most descriptions are referred to a one-handed grasp directed to the handle, sometimes involving the fingers («la tazza, lo stesso, solamente per il manico, con un dito o comunque due dita»). In two cases, the cup is likened to the jug («la tazza girata di qua, uguale alla brocca, cioè farei in modo di comunque avere l'indice dentro il cerchiettino, il pollice sopra e il medio sotto»). The most striking difference with the previous table is that there are seven descriptions that explicitly mention the left hand, five of which are produced by right-handed informants («sempre usando, sì, penso la mano sinistra, perché il manico è verso sinistra»). Only in a few cases is the effector of the grasp the right hand. In such rare cases, it seems that hand dominance predominates over spatial compatibility («dal manico, con l'indice destro e le altre dita che la sostengono»).

The target of the grasp is, in most cases, the handle, usually denoted by meronyms, but sometimes indicated also by spatial terms («questa la prenderei dalla parte sinistra, con due dita magari»). Otherwise, different parts of the object may constitute the target of the grasp and they are usually indicated with words pertaining to the spatial or perceptive domain; for example, participants may describe a grasp directed to the opposite side with respect to the handle («ok, tazzina girata così, io la prenderei con la destra, dalla parte quella grossa»), or from the above («la prenderei probabilmente dall'alto»).

Only a few informants described a two-handed grasp, but in most cases the description also contains a reference to the handle («con tutte e due le mani, una laterale, una per il manico»).

**TEA-CUP (OVERTURNED, RIGHTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (10), *mani* (2).<br>SPA: *destra* (1).<br>QUA: *due* (4), *tre* (2), *cinque* (1), *tutta* (2), *tutte* (2), *sola* (1).<br>MER: *dita* (8), *dito* (1), *pollice* (2), *indice* (1), *medio* (1), *palmo* (3). | ENT: *tazza* (6), *tazzina* (1), *corpo* (1).<br>SPA: *alto* (2), *verso* (1), *base* (1), *fondo* (5), *interno* (1), *lati* (1), *parte* (6), *sotto* (2), *inferiore* (2), *sopra* (3), *intorno* (1).<br>MER: *manico* (4).<br>SPP: *bicchiere* (3), *tazzina* (2). |
| L-H | HND: *mano* (2).<br>SPA: *sinistra* (1).<br>QUA: *tutte* (1).<br>MER: *dita* (1), *pollice* (1), *indice* (1). | ENT: *corpo* (1).<br>SPA: *lato* (1), *basso* (1).<br>MER: *manico* (1).<br>SPP: *pallina* (1), *tazzina* (2). |

When the tea-cup is overturned, it appears that the references to the effector do not change very much, with comparison to the two upright tea-cups: both references to the hand and to the hands' parts are present. But what we notice is a difference in target-related words. Now, the references to the handle are very few («probabilmente manico, una mano, quindi due dita o tre»).

On the contrary, words denoting the upper side of the object (i.e. the base of the overturned cup) in spatial terms increase their frequency; in particular, *fondo*, "bottom" («con tutta la mano, da sopra, e quindi dal fondo della tazza»). The references to the object as a whole are more numerous («con la mano, prenderei la tazza»).

Also, comparison with other objects-stimuli is meaningful. The overturned cup is likened to the glass and the small overturned coffee cup: both of these objects are significant, because the glass is totally lacking any affording part, whereas the coffee cup is overturned (as is the tea-cup) and in most cases its handle is not considered a good target for the grasp (it has only three mentions).

**TEA-CUP (OVERTURNED, LEFTWARD-ORIENTED)**

| | Effector-related words | Target-related words |
|---|---|---|
| R-H | HND: *mano* (15), *mani* (4).<br>SPA: *destra* (5), *sinistra* (2).<br>QUA: *piena* (1), *due* (3), *tutta* (2), *tutte* (2).<br>MER: *pollice* (1), *indice* (1), *dita* (3), *palmo* (2). | ENT: *tazza* (3), *corpo* (2), *oggetto* (1).<br>SPA: *alto* (2), *verso* (1), *basso* (1), *base* (1), *esterno* (1), *fondo* (4), *lato* (1), *parte* (2), *sotto* (1), *inferiore* (1), *superiore* (1), *sopra* (4), *laterale* (1), *lateralmente* (1), *intorno* (1).<br>MER: *manico* (4). |
| L-H | HND: *mano* (2).<br>SPA: *destra* (2).<br>QUA: *due* (1).<br>MER: *dito* (1), *dita* (3). | ENT: *tazza* (1).<br>SPA: *alto* (1), *verso* (1), *lato* (1), *sopra* (1).<br>MER: *manico* (3), *occhiello* (1). |

For the leftward-oriented, overturned tea-cup, we observe again a high number of mentions of the hand, notably of the right hand (both within the left-handed and the right-handed group). The handle is only in rare cases considered a good target of the grasp («questa la prenderei sempre mettendo un dito nell'occhiello»); spatial expressions related to the bottom of the cup, or to its upper part, are more frequent («dall'alto, con tutta la mano e facendo toccare il palmo con il fondo della tazzina»; «con una mano, da sopra»). Sometimes, the side of the object is also preferred to the handle («allora questa la prenderei con la mano destra, lateralmente»).

#### 4.3.3.2 Artefacts with affording parts (with different orientation): general considerations

In Table 4.3, the results of the analyses of effector- and target-related words extracted for the eight artefacts with affording parts (presented with different orientation) are compared. In order to allow a more clear comparison among the different stimuli, data from the leftward and the rightward orientation condition are merged together.

| | EFFECTOR | | | | | TARGET | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | hnd | mer | spa | qua | obp | ent | mer | spa | per | spp |
| Hairdryer | **35** | 17 | 9 | 14 | - | 6 | **47** | 23 | 4 | - |
| Jug | **42** | 15 | 21 | 16 | - | 4 | **47** | 25 | - | - |
| Ladle | 36 | **40** | 18 | 9 | - | 6 | **51** | 40 | 1 | 3 |
| Microphone | **34** | 16 | 16 | 16 | - | 8 | 30 | **32** | 5 | 1 |
| Pacifier | 19 | **38** | 4 | 26 | 1 | 3 | **36** | 30 | 3 | - |
| Sword | **35** | 16 | 10 | 13 | - | 3 | **55** | 18 | - | - |
| Tea-cup | 24 | **60** | 14 | 34 | - | 7 | **43** | 35 | 1 | 2 |
| Tea-cup (ov.) | **35** | 30 | 11 | 22 | - | 16 | 13 | **54** | - | 8 |
| Tot. | **260** | 232 | 103 | 150 | 1 | 53 | **322** | 257 | 14 | 14 |
| % | **34.9** | 31.1 | 13.8 | 20.1 | 0.1 | 8 | **48.8** | 39 | 2.1 | 2.1 |

Table 4.3. Classification of effector-related (tot. 746) and target-related words (tot. 660) provided for the eight artefacts with affording parts, presented with different orientation. Data from the leftward and the rightward orientation condition are merged.

If we consider the frequency data of effector-related words extracted for this group of artefacts, we notice that in most descriptions the reference to the hand(s) is the most frequent one, except for the ladle, the pacifier, and the upright tea-cup, for which participants produced a higher number of words denoting the fingers or other hand's parts. This is not by chance: these three objects-stimuli are those presenting the smallest handle, with comparison to the other objects (such as the microphone or the jug), affording a grip performed with only two or three fingers (indeed, expressions of quantity are also particularly frequent for the pacifier and the tea-cup).

For what regards the target-related words, they mostly denote objects' meronyms. Their frequency values usually exceed (or at least are almost equivalent to) the frequency of words pertaining to the spatial domain (especially for the sword, for which the handle is particularly salient because it allows an agent to avoid touching the blade). In the vast majority of cases, these meronyms denote objects' affording parts, i.e. the parts explicitly designed for the grasp. The most frequent lemmas are those denoting a generic handle (*manico*, *maniglia*, *impugnatura*), which occur 288 times and constitute 89.4% of meronyms. However, only in the case of the overturned tea-cup are spatial relations more frequent than meronymic expressions. This is because the object presented is overturned: for this reason, its handle is not judged to be a probable target of the grasp and thus is rarely mentioned by informants,

who in most cases described a simple and undifferentiated power grasp, directed either to the upper part of the overturned cup or to its side.

## 4.4 Humans

We now turn to the analysis conducted on the three stimuli representing human kinds (the girl, the man, and the little baby).

### 4.4.1 Humans: detailed analysis

**BABY**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (1), *mani* (18).<br>QUA: *due* (10), *tutte* (5), *entrambe* (6).<br>MER: *dita* (2), *pollice* (1), *pollici* (1), *palmi* (1).<br>HOL: *braccia* (4). | ENT: *bambino* (1), *bimbo* (1).<br>SPA: *sotto* (18), *zona* (1), *davanti* (1), *dietro* (1), *centrale* (1).<br>MER: *ascelle* (9), *braccia* (10), *fianchi* (2), *girovita* (1), *gambe* (1), *pancia* (2), *testa* (1), *schiena* (1), *vita* (2), *spalle* (1), *torace* (1), *tronco* (1), *mani* (1), *cosce* (1), *dorso* (1). |

In most of the descriptions provided for the baby, the grasp indicated is performed with two hands, and the baby is grasped under his armpits or under the arms («mettendo entrambe le mani sotto le ascelle del bimbo»; «con i palmi e le dita per contenere, per evitare che cada, da sotto le braccia»); however, many other body parts are named with relation to the target of the grasp («con due mani, al girovita»; «mettendo le mani sotto la pancia del bambino»; «un bambino, allora si può prendere tirandolo per le mani»).

For what regards the effector, most informants opted for a bimanual grasp, but in a few cases we find the arms («con le braccia, penso che lo prenderei»). We notice a high number of quantity expressions: this is mostly due to the fact that informants often specified that they would have grasped the child «con tutte e due le mani», "with both the two hands".

**GIRL**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (3), *mani* (3).<br>QUA: *due* (2). | SPA: *interno* (1), *sotto* (2).<br>MER: *ascelle* (3), *braccio* (8), *braccia* (1), *fianco* (1), *fianchi* (1), *gomito* (3), *spalla* (3), *spalle* (1), *tronco* (1), *vita* (3), *capelli* (1), *mano* (3).<br>AEN: *borsa* (1), *sciarpa* (1), *vestiti* (1). |

For the standing woman, the references to the effector of the grasp are very few.

The most frequent description expresses a grasp (probably a one-handed grasp, even when not explicitly stated) directed to the girl's arm («con una mano, penso dal braccio»). However, there are other possible targets always constituted by body parts, such as the hand, the waist, and many more («la potrei afferrare per il gomito, visto che ce l'ha messo a angolo»; «dalla vita, con due mani»; «probabilmente tenderei a metterle le mani sui fianchi»; «per i capelli»).

Clothes and accessories are rarely chosen as a target of the grasp («prendendo i vestiti, cioè aggrappandomi ai vestiti»; «andrei per la sciarpa»). Clothes were not considered as parts (meronyms) of the girl, but as associated entities. For a sort of metonymic process, the target of the grasp shifts from the girl herself to the garments and accessories that she wears. However, clothes and accessories in some sense are parts of the visual stimulus provided to the informants; therefore, words such as *vestiti*, "clothes", *sciarpa*, "scarf" precisely denote the part to which the grasp is directed (i.e. the target of the grasp).

**MAN**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (3), *mani* (3).<br>QUA: *due* (1), *quattro* (1), *tutte* (1).<br>MER: *pollice* (1), *dita* (1).<br>HOL: *braccia* (1). | ENT: *corpo* (1).<br>SPA: *intorno* (2), *dietro* (1), *bassa* (1), *sotto* (1), *parte* (3).<br>MER: *dorso* (1), *ascelle* (2), *braccio* (8), *braccia* (1), *busto* (1), *gomito* (2), *polso* (2), *pugno* (1), *mano* (3), *schiena* (1), *spalla* (2), *vita* (1), *ginocchio* (1), *petto* (1).<br>AEN: *maglia* (1), *maglietta* (3), *camicia* (1). |

The predominant description provided for the grasp of the running man is a one-handed grasp directed to one of his body parts, in most cases his arm («afferrerei un braccio con la

mano»), otherwise his elbow («per la parte del gomito»), or his hand («per una mano»), but there is a variety of body parts mentioned by informants. We can reasonably assume that in most cases only one hand is involved in the action, but only in three cases did the informants explicitly mention the hand as the effector of the action, whereas a two-handed grasp («subito le braccia, con le mani») is mentioned three times.

In a few cases, participants also described a grasp directed towards his clothes («mi avvicinerei e prenderei la maglietta»; «lo prenderei forse per la camicia, sì»).

It is worth noting that an informant answered that he would have grasped the man with two arms, one under the back and the other under the knees («con le braccia, da una parte prendo la parte delle ginocchia, dall'altra della schiena»). As already observed (cf. §4.2), we can reasonably assume that the hands, too, and not only the arms, are involved as effectors of the grasp. The entity is very large; therefore, as already noticed for the box, it also requires the use of the arms.

### 4.4.2 Humans: general considerations

Looking at Table 4.4, we can compare the data extracted and classified for the category of human kinds.

| | EFFECTOR | | | | TARGET | | | |
|---|---|---|---|---|---|---|---|---|
| | hnd | mer | hol | qua | ent | mer | spa | aen |
| Baby | 19 | 5 | 4 | **21** | 2 | **35** | 22 | - |
| Girl | **6** | - | - | 2 | - | **29** | 3 | 3 |
| Man | **6** | 2 | 1 | 3 | 1 | **27** | 8 | 5 |
| Tot. | **31** | 7 | 5 | 26 | 3 | **91** | 33 | 8 |
| % | **44.9** | 10.1 | 7.3 | 37.7 | 2.2 | **67.4** | 24.5 | 5.9 |

Table 4.4. Classification of effector-related (tot. 69) and target-related words (tot. 135) provided for the three human kinds.

For the stimuli of the two adults, we observe a strong tendency to name a body part as the target of the grasp, whereas clothes and accessories are chosen as possible targets only, respectively, in three and six cases. Body parts well-suited for the grasp are usually the most protruding ones, i.e. the arms and the hands.

As for the baby, we observe a different tendency: protruding body parts such as the arms and the hands are rarely named, probably because they are considered too fragile; most grasps, described as bimanual grasps, are directed to the sides of the baby. In some cases, the arms are involved in the action to provide further support for the little child.

No informant described a precision grip: the reference is always to a power grasp, performed with the whole hand. In the rare cases in which the fingers are mentioned, they are never indicated as the effector of a pinch grasp.[41]

## 4.5 Natural kinds

This paragraph presents the results of the analysis conducted on the four natural kinds (the mandarin, the apple, the banana, and the stone).

### 4.5.1 Natural kinds: detailed analysis

#### MANDARIN

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (26); *pugno* (1).<br>SPA: *destra* (1), *dentro* (1).<br>QUA: *due* (1), *tre* (1), *sola* (3), *solo* (2), *piena* (2), *tutte* (1), *tutto* (1), *tutta* (4).<br>MER: *palmo* (2), *palmi* (1), *dita* (5), *pollice* (2), *indice* (2). | ENT: *mandarino* (4), *oggetto* (2), *frutto* (1).<br>SPA: *bordo* (1), *lato* (1), *lati* (1), *sopra* (3), *intorno* (3), *attorno* (1).<br>QUA: *tutto* (1).<br>SPP: *palla* (1), *pallina* (6). |

In most descriptions, the mandarin is grasped with only one hand; therefore, the references to the hand (in particular *mano*) are the most numerous among the effector-related words («come la pallina, semplicemente con la mano piena, raccolta»; «questo con tutta la mano, col palmo, direttamente»). Only five participants named the fingers and not simply the hand («lo prendo da sopra afferrandolo con tutte le dita»; «questo con tre dita, dai lati»).

As to target-related words, we notice that sometimes the whole object is explicitly indicated as the target of the grasp («con la mano ad avvolgere il mandarino»; «con una mano,

[41] In general, the mention of the fingers is not strictly related to the description of a pinch grasp: on this topic, cf. the discussion in §4.7.1.1.

avvolgendo la mano attorno al frutto»). However, in most cases, the target is indicated in terms of spatial relations («con una mano sola, da sopra, cioè lo sollevo da sopra praticamente»). It is worth noting that seven informants explicitly referred to the similarity between this stimulus and a tennis ball (e.g. «come la pallina da tennis, con la mano tutta intorno»; «lo afferrerei con la mano come si fa con una pallina»).

**APPLE**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (17), *pugno* (1).<br>SPA: *destra* (1), *sinistra* (1).<br>QUA: *due* (2), *tre* (1), *sola* (2), *piena* (1), *tutte* (1), *tutto* (1), *tutta* (3).<br>MER: *palmo* (3), *dita* (7), *pollice* (3), *indice* (2), *dito* (1). | MER: *gambetto* (1), *picciolo* (4).<br>SPA: *parte* (2), *laterale* (1), *destra* (1), *intorno* (1), *sopra* (2).<br>QUA: *tutta* (1).<br>SPP: *pallina* (2), *mandarino* (1). |

By looking at effector-related words elicited by the apple, we can easily observe that they all indicate a grasp performed by a single hand and, in particular, a whole hand, as many adjectives suggest. Accordingly, the target of the grasp described in the vast majority of cases is just the whole object («anche questa con tutta la mano»; «questa anche con tutto il palmo e le dita intorno»; «con una mano, semplicemente afferrandola tutta»). Only in five cases did informants name the stalk («dal picciolo, con due dita»), whereas some of them also referred to which part of the apple would be in contact with the hand (describing a grasp either from the above or laterally). It is interesting to observe that in three cases the grasp directed to the apple is explicitly said to be similar to that of the tennis ball and of the mandarin.

**BANANA**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (18); *mani* (2).<br>SPA: *destra* (1), *mezzo* (1), *interno* (1).<br>QUA: *due* (2), *tre* (2), *quattro* (1), *sola* (3), *tutta* (2).<br>MER: *palmo* (3), *dita* (6), *pollice* (2), *indice* (1). | ENT: *oggetto* (1).<br>MER: *gambo* (1), *picciolo* (3).<br>SPA: *alto* (1), *centro* (1), *là* (1), *qua* (1), *intorno* (1), *circonferenza* (1), *mezzo* (1), *metà* (1), *lato* (1), *sopra* (2), *punte* (1), *punta* (1).<br>QUA: *tutta* (1).<br>SPP: *coltello* (1), *mandarino* (1). |

For the banana, again, most informants described a general one-handed grasp directed to the whole object («questa con tutta la mano»; «banana, la impugno tutta»), whereas some of them referred to the fingers, and only a few named two hands. As to the target of the grasp, it is usually denoted by spatial terms («con due dita che stringono la circonferenza»). Only in four cases is the stalk of the banana explicitly mentioned («dal picciolo, con tre dita»).

Again, we find an explicit comparison with the mandarin («da stessa cosa che per il mandarino, cioè la prenderei con una sola mano»), as well as one with *coltello*, "knife" («da banana, la impugnerei come impugnerei un coltello, quindi la prenderei che me la faccio passare in mezzo al palmo e poi la stringerei»).

**STONE**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (23). | ENT: *oggetto* (1), *pietra* (2), *sasso* (1). |
| OBP: *piede* (1). | SPA: *alto* (1), *parte* (1), *sotto* (1), *sopra* (3), *intorno* (1), *superficie* (1), *attorno* (1). |
| SPA: *destra* (1), *dentro* (1). | QUA: *tutto* (2). |
| QUA: *sola* (4), *piena* (1), *intera* (1), *tutto* (2), *tutta* (7), *tutte* (1). | SPP: *pallina* (4), *mandarino* (1), *arancia* (1). |
| MER: *palmo* (3), *dita* (6), *pollice* (1), *dorso* (1), *polpastrelli* (1). | PER: *ampia* (1). |

In most descriptions, the stone is grasped with a whole hand, as indicated by the high frequency of words expressing quantity («con tutta la mano»; «questo con tutta la mano, direttamente»). The fingers are rarely mentioned, and always with reference to a power grasp («la pietra la prenderei sicuramente solo con la mano aperta, e le dita a chiusura»). Only in one case is the foot indicated as the effector involved in the action, instead of the hand («mah… con un piede»).

Since the stone has no meronyms, the reference to the target of the grasp is made through spatial or perceptive expressions («questo, credo forse appoggiando sopra il pollice e le altre dita sotto»; «questo con tutto il palmo, le dita che si chiudono attorno»). Otherwise, the target is explicitly constituted by the whole object («una mano intorno a tutto all'oggetto»), or it is left unexpressed.

It is worth noting that in six cases participants stated the similarity with the tennis ball, the mandarin, and an orange («questa la prenderei dall'alto sempre come si afferra una

pallina»; «il sasso lo prenderei a mano piena, cioè proprio come la pallina da tennis e il mandarino»).

### 4.5.2 Natural kinds: general considerations

We can now compare the data related to the four natural kinds and gather them together.

| | | EFFECTOR | | | | TARGET | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | hnd | spa | qua | mer | obp | ent | mer | spa | qua | per | spp |
| Mandarin | **27** | 2 | 15 | 12 | - | 7 | - | **10** | 1 | - | 7 |
| Apple | **18** | 2 | 11 | 16 | - | - | 5 | **7** | 1 | - | 3 |
| Banana | **20** | 3 | 10 | 12 | - | 1 | 4 | **13** | 1 | - | 2 |
| Stone | **23** | 2 | 16 | 12 | 1 | 4 | - | **9** | 2 | 1 | 6 |
| Tot. | **88** | 9 | 52 | 52 | 1 | 12 | 9 | **39** | 5 | 1 | 18 |
| % | **43.6** | 4.5 | 25.7 | 25.7 | 0.5 | 14.3 | 10.7 | **46.4** | 6 | 1.2 | 21.4 |

Table 4.5. Classification of effector-related (tot. 202) and target-related words (tot. 84) provided for the four natural kinds.

By looking at effector-related words extracted from the descriptions of grasp of a natural kind, it seems that most answers refer to a one-handed grasp performed with the whole hand. *Mano*, "hand", is the word most frequently named for all stimuli (23 times for the stone; 26 for the mandarin; 18 for the banana; 17 for the apple). The second word most frequently named with reference to the effector is *dita*, "fingers" (six times for the stone and the banana; five for the mandarin; seven for the apple).

In the previous chapter (cf. §3.4), we observed that informants rarely indicated the target of the grasp of a natural kind, compared to other objects' categories. Now, we can also add that the target of the grasp is generally described with words pertaining to the visuo-spatial domain, even when the object-stimulus presents distinguishable parts (the stalks of the apple and the banana). Such parts are rarely mentioned (only nine occurrences, in the 60 descriptions provided for the apple and the banana); therefore, they seem not to be considered as a good target for the grasp, probably because stalks do not play an important role in the actions they are usually involved in. For instance, we rarely hold them in our fingers while we are eating or peeling fruits and, when we have to take a banana or an apple,

in order to move it from one place to another, or to put it in the fridge or in a can, we usually prefer a simpler and faster power grasp.

Sometimes, the descriptions of the target of the grasp also contain a reference to the object as a whole: this seems to happen especially for objects that lack specific parts, i.e. the stone and the mandarine. For these objects, explicit references to other objects-stimuli affording an undifferentiated one-handed grasp (such as the tennis ball) are also found.

## 4.6 Substances and aggregates

This section reports the results of the analysis conducted on the last category of objects-stimuli, the one collecting substances and aggregates (the water, the flour, the sand, and the pumpkin seeds).

### 4.6.1 Substances and aggregates: detailed analysis

**WATER**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (7), *mani* (20).<br>INS: *bicchiere* (1), *recipiente* (1).<br>SPA: *perpendicolare* (1), *verso* (1), *alto* (1).<br>QUA: *due* (7), *entrambe* (2), *tutte* (3).<br>MER: *palmo* (1), *palmi* (1), *dita* (2).<br>PER: *concava* (1).<br>SPP: *conca* (6), *coppa* (5), *conchetta* (2), *contenitore* (3), *ciotola* (1), *piscinetta* (1), *utensile* (1). | ENT: *acqua* (3).<br>SPA: *sotto* (5).<br>SPP: *sabbia* (2). |

The most frequent grasp type that participants referred for the water is with two cupped hands («la prenderei con due mani, quindi chiudendo le due mani a conchetta in modo da poter mantenerla dentro, altrimenti con una mano sola non ci si riesce, perché scappa»; «metterei le due mani a mo' di piscinetta, come si dice, di contenitore»; «userei le mani a ciotola, un po' come ho fatto con la sabbia in un certo senso»). The equivalent action, performed with only one cupped hand, is chosen only by seven informants («usando la mano

a coppa, come un utensile»; «con una mano, facendo a mo' di coppa, cercando di trattenere acqua»).

In two cases, the grasp described is performed by using a container («ci metterei un recipiente sotto, di qualunque dimensione, a seconda di quanta me ne serve»; «con un bicchiere, fondamentalmente»). Finally, one informant said, with great uncertainty and hesitation, that she would have grasped the water with her fingers, but this action does not lead, obviously, to holding the substance («con le dita»).

The only spatial term referred to the water is *sotto*, "under", because in the picture the water flows from a faucet («ci infilo le mani sotto, ma visto che è liquida è difficile che riesca a fermarla»). It is also worth noting that, in two cases, the grasp of the water is likened to the grasp of the sand («mi aiuterei con le mani, come faccio per la sabbia»).

**FLOUR**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (14), *mani* (10), *pugnetto* (2), *pugno* (2). | ENT: *farina* (3). |
| INS: *bicchiere* (1), *tazza* (1). | SPA: *intorno* (1). |
| SPA: *destra* (1). | QUA: *po'* (1), *più* (1), *poca* (1), |
| QUA: *due* (6), *entrambe* (3), *sola* (1), *piena* (1), *intera* (1), *tutte* (1). | *quantità* (1), *mucchietto* (1), *manciata* |
| MER: *palmo* (1), *dita* (5). | (3). |
| SPP: *coppa* (3), *conchetta* (1), *contenitore* (1), *mestolo* (1), *cucchiaio* (1). | SPP: *sabbia* (1), *acqua* (1). |

For the flour, the type of grasp most frequently described is with the whole hand or the fist («a mano piena»; «mettendo appunto la mano intera e chiudendo le dita intorno alla farina»). Only in one case did a participant indicate the fingers referring to a pinch grip («se ne devo prendere poca, anche con due dita»).

More interestingly, there are many participants that described the shape of the hand by likening it to a container («come l'acqua, formerei una conchetta con le mani quindi aperte e man mano le andrei a chiudere per prendere la farina»; «usando una mano come una coppa per raccogliere la farina»; «la farina, sempre tenendo la mano semichiusa la solleverei eeh come se utilizzassi appunto un cucchiaio, un mestolo). In such descriptions, the effector of the grasp is still a body part (either one or two hands); therefore, participants stated a similarity between their hands and a container. But, surprisingly, two informants also named

a real container as effector of the grasp («la farina la prenderei con un bicchiere»; «mi aiuterei con una tazza, per raccoglierla»).

These two words related to the effector have been grouped together into the category instrument, because they actually denote the entity with which the flour comes in contact. It is interesting to note that the body effector that controls the instrument, i.e. the hand holding the glass and the cup, is never explicitly mentioned.

**SAND**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (17), *mani* (10), *pugnetto* (1), *pugno* (3).<br>SPA: *sinistra* (1), *destra* (1), *parte* (1), *sotto* (1), *perpendicolare* (1).<br>QUA: *due* (3), *tre* (1), *quattro* (1), *entrambe* (3), *solo* (1), *piena* (1), *intero* (1), *tutta* (2).<br>MER: *palmo* (2), *dita* (4).<br>SPP: *coppa* (1), *conca* (1), *conchetta* (1), *braccio* (1), *contenitore* (1), *mestolo* (1), *utensile* (1), *cucchiaio* (1). | ENT: *sabbia* (8).<br>SPA: *attorno* (1), *dentro* (1), *interno* (1), *alto* (1), *basso* (1).<br>QUA: *mucchi* (1), *quantità* (3), *pugno* (2), *più* (3), *manciata* (3).<br>MER: *granelli* (1). |

According to the descriptions collected from the informants, the sand can be grasped either with one or two hands («con tutta la mano, la raccoglierei solo con una mano»; «cercherei diciamo di afferrarla con un pugno»; «con due mani»).

Again, we observe a high number of references to containers or instruments (even to the arm of an excavator), in order to describe the cup-shaped hands («farei una conca tipo con la mano»; «andrei a scavare con la parte sotto diciamo della mano, a mo' di cucchiaio»; «con tutta la mano usandola come un mestolo, insomma, un utensile»).

In a few cases, the target of the grasp is indicated in spatial terms («ci affonderei la mano dentro»; «chiuderei appunto le dita attorno alla sabbia, stringendo il più possibile»); otherwise, only the sand itself is named («aprirei la mano, raccoglierei la sabbia»).

It is interesting to note the high frequency of words expressing the quantity of sand that would be grasped («farei una conca tipo con la mano, sì, cercherei di prenderla… di prenderne di più»; «la prenderei con due mani, per prenderne la maggior quantità possibile»; «dovessi effettivamente prenderne una manciata, penso delle mani, più che altro, sì»).

**PUMPKIN SEEDS**

| Effector-related words | Target-related words |
|---|---|
| HND: *mano* (10), *mani* (1), *pugno* (1).<br>SPA: *sinistra* (1), *destra* (1).<br>QUA: *due* (10), *tutta* (2), *tutte* (1).<br>MER: *palmo* (1), *pollice* (7), *punta* (2), *indice* (7), *dita* (17), *unghie* (1). | ENT: *semi* (1).<br>SPA: *sotto* (1), *sopra* (1), *attorno* (1).<br>QUA: *tutti* (1), *uno ad uno* (2), *uno alla volta* (3), *uno per volta* (2), *singolarmente* (1), *più* (1), *manciata* (4).<br>MER: *semino* (2).<br>SPP: *sabbia* (1). |

Most participants, in front of the picture representing a mound of pumpkin seeds, described a grasp directed to a single seed. Moreover, because of the very small size of such objects, they often named the fingers (typically the thumb and the index), thus denoting a pinch grip («pollice e indice, afferro il semino»; «questi con due dita»; «con la punta delle dita»). However, there are also a few mentions of a grasp performed with the whole hand, in a fist-like manner («con tutta la mano»), whereas only in one case do we find a reference to a bimanual grasp («li raccoglierei con tutte e due le mani»).

For what regards the target of the grasp, we should consider that, in the picture, the pumpkin seeds are presented in a compact mound, similar to that of the sand and of the flour (typical mass entities), but still, discrete physical objects can be distinguished in the image. It seems that, in this grasp description task, the pumpkin seeds presented in a mound are mostly considered as an aggregation of individuals, each individual being a possible target of the grasp («questi li raccoglierei uno ad uno, quindi sempre con le dita»). This is probably due to the most usual mode of interaction that participants have with the seeds, which are commonly eaten one by one.[42] This is also the reason why references to hand's parts (especially fingers) are so high. The single seeds are rarely mentioned in an explicit form (such as *semino*, "little seed"), but mostly by means of lexical expressions denoting quantity, such as "one by one" or "singularly" («raccogliendoli o uno ad uno con due dita»; «prenderne uno singolarmente, lo prenderei con pollice e indice della mano sinistra»).

However, as already said, there is also a competing grasp description reflecting a different conceptualisation of the object-stimulus. In a few answers, the seeds are grasped with the

[42] The mode of interaction with the relevant entity, as well as the possibility of distinguishing its constituent elements, are also important in determining a noun's classification; cf. Wierzbicka (1988).

whole hand («con tutta la mano, cercando di prenderne il più possibile, come per raccogliere della sabbia»); the last example is also significant, because the informant stated a similarity between the pumpkin seeds and the sand. In such descriptions, it seems that the mass interpretation, fostered by the presence of a "mound" of seeds, overcomes the individuation of the single objects aggregated.

It is worth noting that sometimes the two different competing strategies of grasp are evoked: «senza prendere una manciata, uno alla volta con due dita»; «questi mi devo aiutare con le dita, li prenderei uno per volta, perché non si possono prendere tutti insieme».

Looking at the very few occurrences of the lemma *seme*/*semino* ("seed"/"little seed"), when the reference to the entity is plural, participants seem to conceptualise the seeds as an aggregate and describe a power grasp («tutta la mano, chiudendo le dita attorno ai vari semi»); instead, when they refer to the single seed, they indicate a single "particle" of the mound and describe a pinch grasp. This is the reason why *semino* was included in the category of meronyms (cf. Cruse 1986), whereas *semi* better denotes the entity represented in the picture, i.e. the seeds aggregated in a mound.

### 4.6.2 Substances and aggregates: general considerations

Table 4.6 collects all data extracted and classified for the category of substances and aggregates.

| | EFFECTOR | | | | | | | TARGET | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | hnd | spa | qua | mer | spp | per | ins | ent | mer | spa | qua | spp |
| Flour | **28** | 1 | 13 | 6 | 7 | - | 2 | 3 | - | 1 | **8** | 2 |
| Sand | **31** | 5 | 13 | 6 | 8 | - | - | 8 | 1 | 5 | **12** | - |
| Pumpkin seeds | 12 | 2 | 13 | **35** | - | - | - | 1 | 2 | 3 | **14** | 1 |
| Water | **27** | 3 | 12 | 4 | 19 | 1 | 2 | 3 | - | **5** | - | 2 |
| Tot. | **98** | 11 | 51 | 51 | 34 | 1 | 4 | 15 | 3 | 14 | **34** | 5 |
| % | **39.2** | 4.4 | 20.4 | 20.4 | 13.6 | 0.4 | 1.6 | 21.1 | 4.2 | 19.7 | **47.9** | 7.1 |

Table 4.6. Classification of effector-related (tot. 250) and target-related words (tot. 71) provided for the four substances and aggregates.

Considering effector-related words, we observe that the grasp of the sand and the flour is mostly described as involving one or two hands, whereas for the pumpkin seeds the fingers

are most likely to be named. However, the most striking evidence is that there is a very high number of references to containers, either as the real instrument with which the entity is grasped (four cases), or evoked to describe the shape of the hand (this argument will be explored further in §4.7.1.3).

It is not by chance that containers are named especially for the water, the flour and the sand: these are true mass entities, both from a linguistic and a conceptual point of view. Their component parts are continuous and not clearly distinguishable from one another. The action of grasp implies a form of control over the grasped entity. In the case of substances and mass entities, this control cannot be obtained with the hands, as participants sometimes explicitly stated (e.g. for the water: «con una sola mano non ci si riesce, perché scappa»). Instead, for the pumpkin seeds, containers and instruments are never mentioned, because an agent can grasp the single seeds with a pinch grip, or also a fistful of seeds, with the whole hand (and, in this case, a certain quantity of seeds will remain in his hand, whereas for the water this is almost impossible).

Therefore, it seems that the number of explicit mentions of containers outlines a distinction into this class of objects-stimuli. The water is a liquid substance that cannot be grasped with the hands unless they are cupped, i.e. shaped as (and linguistically assimilated to) a container; otherwise, a real container, such as a glass, must be used. The flour and the sand can be grasped with a pinch grasp or with a power grasp (a fistful of flour or of sand), but still hands are often assimilated to containers; moreover, for the flour (whose particles are smaller than those of the sand), a real container is indicated as the instrument with which the action can be performed (as for the water). On the other hand, the mound of pumpkin seeds is mostly considered as an aggregate of individuals: containers are never mentioned, whereas the fingers are very often indicated as effectors of the grasp.

A similar distinction among these objects-stimuli can be drawn by considering target-related words. It seems that the expression of the quantity of the entity that can be grasped is quite frequent for the flour and the sand, and especially for the pumpkin seeds, whereas it is never found in the descriptions provided for the water.

This is hardly surprising: again, the pumpkin seeds and the water are at the two opposite poles of a sort of hierarchy. For the seeds, we have observed that participants gave two different kinds of descriptions that reflect a conceptualisation of the mound either as a unit, or as an aggregate of individuals (in which each seed might be a possible target of the grasp). The sand, the flour and the water could be associated to words expressing their quantity, but,

in fact, this happens quite rarely, and never for the water. This is not because the water cannot be quantified, when obviously it can; rather, it is probably because the quantification of the amount of water that one can grasp is not relevant from a cognitive point of view.

Again, we should consider not only the constituency of the entity, but also the way in which we generally use the water flowing from a faucet. In very few occasions we have to pay attention to the exact quantity of the substance we are taking; therefore, quantity is not an aspect of the water that we usually consider salient. On the contrary, we often eat pumpkin seeds one by one. The sand and the flour are in between: since they are granular aggregates made of solid particles, a human is able to grasp a certain quantity of sand or flour, even by using only his hands (whereas for the water this is not possible). However, it would still be difficult to determine this quantity with linguistic means, unless with vague quantifying expressions (such as "a lot"), whereas seeds, in Italian, are denoted by a count noun and can be quantified with number words (cf. also §2.4.3.2).

Many studies remark that the syntax of nouns referring to aggregates is systematically related to differences in how people perceive and interact with such aggregates (e.g. Middleton et al. 2004): the results emerging from this study well comply with this statement.

## 4.7 General results and discussion

### 4.7.1 Words related to the effector of the grasp

Table 4.7 summarises the results of the analysis conducted in this chapter for effector-related words (both frequencies and percentages are reported).

It is evident that, within each category of objects-stimuli, most effector-related words simply denote the hand as the effector of the grasp. However, we can make some further remarks on these data and, in particular, there are three aspects that are worth mentioning and will be discussed in turn.

| | | hnd | mer | hol | obp | spa | perc | qua | spp | ins | tot. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Humans | freq. | 31 | 7 | 5 | - | - | - | 26 | - | - | 69 |
| | % | **45** | 10.1 | 7.2 | - | - | - | 37.7 | - | - | |
| Artefacts with AP | freq. | 372 | 273 | 9 | 1 | 119 | - | 204 | 2 | - | 980 |
| | % | **38** | 27.9 | 0.9 | 0.1 | 12.1 | - | 20.8 | 0.2 | - | |
| Artefacts without AP | freq. | 181 | 145 | 9 | - | 23 | - | 166 | - | - | 524 |
| | % | **34.5** | 27.7 | 1.7 | - | 4.4 | - | 31.7 | - | - | |
| Natural kinds | freq. | 88 | 52 | - | 1 | 9 | - | 52 | - | - | 202 |
| | % | **43.6** | 25.7 | - | 0.5 | 4.5 | - | 25.7 | - | - | |
| Substances/ Aggregates | freq. | 98 | 51 | - | - | 11 | 1 | 51 | 34 | 4 | 250 |
| | % | **39.2** | 20.4 | - | - | 4.4 | 0.4 | 20.4 | 13.6 | 1.6 | |
| Tot. | freq. | 770 | 528 | 23 | 2 | 162 | 1 | 499 | 36 | 4 | 2025 |
| | % | **38.1** | 26.1 | 1.1 | 0.1 | 8 | 0.1 | 24.6 | 1.7 | 0.2 | |

Table 4.7. Classification of effector-related words with relation to the different categories of objects-stimuli: frequencies data and percentages (over the total number of effector-related words extracted for each category).

#### 4.7.1.1 The hand and the hand's parts

Much of the research conducted on grasping in diverse fields, such as kinesiology, robotics, artificial intelligence, and rehabilitation accord primary importance to the study of the possible configurations of the hand (e.g. MacKenzie and Iberall 1994). There exist many taxonomies that have been proposed in order to classify grasps, mostly considering differences in the hand's shape. Just to give an example, we may consider the GRASP taxonomy,[43] one of the most comprehensive resources, which is based on a definition of grasp as every static hand posture with which an object can be held securely with one hand (Feix et al. 2009). In order to build this taxonomy, most of the extant classifications developed in different fields from robotics, to developmental medicine, to biomechanics (e.g. Cutkosky and Wright 1989; Kang and Ikeuchi 1992) were analysed and compared to find the maximum number of grasp types. In total, the authors (Feix et al. 2009) found 147 grasp examples in the considered literature sources, among which they identified only 45 different

[43] The taxonomy was developed within the European Union Project GRASP (http://www.csc.kth.se/grasp/).

grasp types. A further classification, based on the grasp definition provided above, revealed only 33 valid grasp types[44] that constitute the taxonomy (Fig. 4.1):

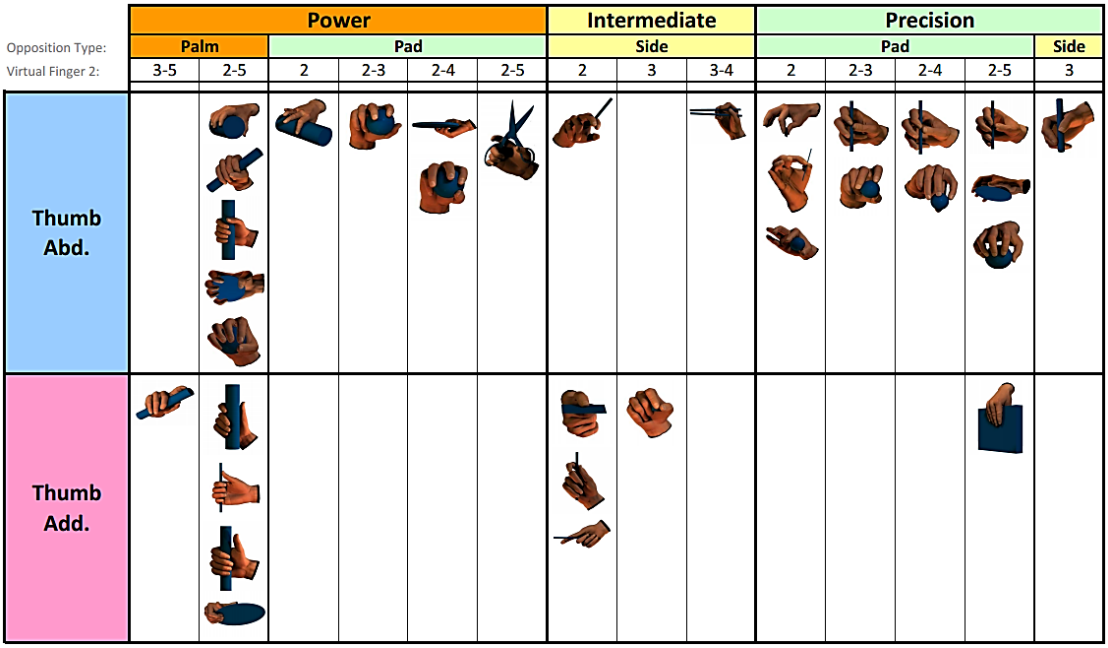



Figure 4.1. The Grasp Taxonomy (retrieved from: http://grasp.xief.net/documents/taxonomy.pdf; *abd.*: abducted, *add.*: adducted).

This is just an example of how rich a taxonomy of grasps might appear. If we now consider the data collected in the present study, we observe that even the most widely accepted distinction, i.e. the one between power and precision grip (note that in Fig. 4.1 also an "intermediate" category is inserted), is scarcely reflected in the linguistic descriptions of grasps, as will be briefly clarified.

Table 4.7 clearly shows that the most frequent words related to the effector found in the transcripts simply denote a hand (or at most two hands). However, there are descriptions that contain words related to hands' parts, and therefore might indicate a somehow more specific grasp type. Words denoting a hand's meronyms constitute 26.1% of the total of words extracted from the references to the effector, and in the detailed analysis provided for the single objects in the previous paragraphs, we have often observed that the

[44] For instance, the authors did not consider as an example of a valid grasp the position of a flat hand under an object, although it is a kind of grasp included in some broader taxonomies.

mention of fingers (the most frequently named hand's parts) is likely to indicate a pinch grip (e.g. for the pencil and the lighter; cf. §4.3.1.1). However, we should observe how these words are distributed in the descriptions provided for the objects-stimuli. Table 4.8 reports the number of descriptions provided by informants (from 0 to 30) that contain a reference to the hand (*mano*, *mani*, *pugno*), compared to the number of descriptions that contain one or more words denoting fingers or their parts (such as *dito*, *dita*, *pollice*, *indice*, *medio*, *mignolo*, *polpastrello*, etc.):

| Object | H | F | Object | H | F | Object | H | F |
|---|---|---|---|---|---|---|---|---|
| pencil | 28 | 23 | apple | 19 | 8 | vase | 19 | 4 |
| pumpkin seeds | 10 | 23 | banana | 19 | 7 | jug (L) | 17 | 4 |
| tea-cup (R) | 10 | 21 | tennis ball | 29 | 6 | girl bag | 12 | 4 |
| pacifier (R) | 11 | 15 | mandarin | 26 | 6 | baby | 19 | 3 |
| lighter | 20 | 14 | stone | 23 | 6 | umbrella | 18 | 3 |
| coffee cup (O) | 19 | 14 | jug (R) | 23 | 6 | sword (L) | 15 | 3 |
| tea-cup (L) | 13 | 13 | hairdryer (R) | 15 | 6 | box | 27 | 2 |
| pacifier (L) | 8 | 13 | flour | 24 | 5 | water | 26 | 2 |
| tea-cup (O, R) | 14 | 11 | sword (R) | 19 | 5 | rubber boat | 15 | 2 |
| glass | 23 | 10 | microphone (R) | 14 | 5 | trolley | 11 | 2 |
| ladle (L) | 20 | 10 | sand | 27 | 4 | backpack | 13 | 1 |
| tea-cup (O, L) | 20 | 9 | chair | 23 | 4 | man | 6 | 1 |
| ladle (R) | 15 | 9 | hairdryer (L) | 20 | 4 | soccer ball | 27 | 0 |
| plate | 23 | 8 | microphone (L) | 19 | 4 | girl | 6 | 0 |

Table 4.8. Number of descriptions containing an explicit reference to a "hand" (H) effector (*mano*, *mani*, *pugno*) and to a "finger" (F) effector.

We are not allowed to consider words such as *mano*, *mani* ("hand", "hands") as necessarily indicating a true power grasp, i.e. a grip performed with the whole hand: their occurrence may just be due to the genericity or under-specification of the answer. Similarly, not all references to the fingers necessarily indicate a precision grip. We may think, for example, of a whole-handed grasp of a tennis ball described as "a grasp with the fingers bent around the object".

However, even if we arbitrarily assumed that informants used these words in order to describe a precision grip, the stimuli for which at least 50% of the informants named fingers or fingers' parts are only the pumpkin seeds, the rightward-oriented tea-cup, the rightward-

oriented pacifier, and the pencil. Therefore, even if there is a tendency to name these effector-related words in describing the grasp of a small or very thin object, this trend is not consistently followed by informants. It is very frequent to find, for the same object, descriptions that probably denote the very same action, but contain different effector-related words. For instance, for the leftward-oriented ladle, we have «dal manico, usando indice e pollice» ("by the handle, using my thumb and my index"), as well as «manico, con la mano» ("handle, with my hand"). It is interesting to note that the ladle is explicitly likened three times to a pen during the interviews, but only nine and ten descriptions (respectively for the right and the left condition) contain one or more explicit mentions of the fingers, whereas for the pencil, which is the object more similar to a pen, there are 23.

Table 4.8 also shows that the descriptions that contain more mentions of the fingers also contain mentions of the hand: only for the pumpkin seeds, the rightward-oriented tea-cup, and the pacifier (in both orientation condition), are the descriptions in which the fingers are named more frequent than those in which the hand is named.

There is also a certain variation within the answers provided by the same participant for different stimuli. For example, a student described a precision grasp for the leftward-oriented pacifier («con le dita, dal manico», "with my fingers, by the handle"), whereas for the rightward-oriented pacifier he named only the right hand and the handle («con la mano destra, sempre dal manico», "with my right hand, by the handle").

Therefore, the objects-stimuli that afford a precision grip (for instance the pacifier and the lighter), for which informants *consistently* named the fingers (either with or without any mention of the hand), are very few. On the contrary, the fingers are named also with relation to objects that clearly require a power grasp, such as the vase (e.g. «chiudendo le dita intorno alla circonferenza e alzandolo»). Probably, the only indicator of a real precision grip is the reference to fingertips (*punta delle dita*); however, the hand meronym *punta* occurs only five times in the entire corpus (in two descriptions provided for the pencil, two for the pumpkin seeds, and one for the overturned coffee cup). Expressions of quantity may be helpful to distinguish between the descriptions referring to a grasp performed with all the fingers and those referring to a grasp performed with only two or three fingers (likely to indicate a power grasp and a precision grasp, respectively). However, as the analysis of the extractions for the single objects-stimuli has shown, expressions of quantity are not only associated with the hand's meronyms: for instance, whereas 23 descriptions provided for the pencil contain a reference to the fingers, only 10 of them also specify the number of fingers involved in the

action (two fingers, three fingers); similarly, 23 descriptions provided for the pumpkin seeds contain a reference to the fingers, but only 11 also indicate the quantity of fingers required by the grasp (two fingers).

A more in-depth study could be conducted in order to find correspondences between the lexical content of the transcripts of the interviews and a taxonomy of grasps, even a very basic one that only distinguishes between a power and a precision grip, but this research goes beyond the scope of this work. Therefore, the first impression gained from this analysis is that even the most commonly accepted distinction in grasp taxonomies, i.e. between power and precision grasp, seems not to be consistently and clearly reflected in the lexical choices of the informants. As already stated, there is some inconsistency in signalling this distinction across the descriptions collected for a single object-stimulus, as well as within the linguistic production of a single participant. Generic expressions such as "with my fingers" or "with my hand", may be referred both to large and small objects that respectively afford a power and a precision grip. Only for two stimuli out of 42, i.e. the soccer ball and the girl, are the fingers never mentioned.

Last, but not least, there is a strong limitation that should be kept in mind: many descriptions do not contain any reference at all to the effector of the grasp, as the analysis of Chapter III has shown (cf. §3.6). In such cases, language provides no cues as to whether the effector of the grasp is the hand as a whole, or specifically the fingertips; therefore, any comparison with a grasp taxonomy could not be established.

#### 4.7.1.2 The visuo-spatial domain

Considering again Table 4.7, we notice that there is only one perceptive adjective produced with reference to the hand, i.e. *concavo* ("concave", for the water stimulus); expressions such as *mano curva*, *mano ricurva*, *mano piatta* ("curved hand", "flat hand", both in their singular and plural form) do not occur in the entire corpus. We can therefore merge data from the perceptive and spatial domain into a single category. Considering the detailed analysis provided for the single objects, we observe that words pertaining to the visuo-spatial domain, in most cases, simply specify the side of the hand indicated as effector of the grasp: almost 90% of the words of this category (145 over 163) is constituted by *destra* or *sinistra* ("right" and "left"). Moreover, spatial references to the hand are mostly produced for artefacts with affording parts, especially those presented with different orientation (cf. Tables

4.2, p. 125, and 4.3, p. 139). In this regard, we can focus on this class of stimuli (cf. §4.3.3.1) and consider how the lemmas *destro* and *sinistro* are distributed in the descriptions provided by the two groups of informants (righthanded vs. lefthanded) for rightward- and leftward-oriented objects-stimuli:

| | Right-handed | | Left-handed | |
|---|---|---|---|---|
| | *destro* | *sinistro* | *destro* | *sinistro* |
| Rightward-oriented | 24 | 1 | 6 | 4 |
| Leftward-oriented | 18 | 34 | 4 | 10 |

Table 4.9. Occurrences of the lemma *destro* ("right") and *sinistro* ("left") with reference to the effector of the grasp in the descriptions provided for the eight artefacts with affording parts, presented with different orientation (16 stimuli).

Most of the occurrences of *destro* and *sinistro* are produced when the hand chosen as the effector of the grasp described coincides with the orientation of the handle (58 cases for the right-handed group, 16 cases for the left-handed group). For the right-handed group, there is a tendency to specify which hand would be the best effector for the target described, especially when the orientation of the handle does not correspond to the dominant hand of the informant (34 vs. 24 cases), as if the informant signalled the anomaly of choosing the non-dominant hand in order to preserve the spatial compatibility between the effector and the target of the grasp (that in most cases is the handle of the object).

#### 4.7.1.3 Instruments and containers

The analysis conducted in this chapter also considers all words denoting a concrete entity evoked by informants in order to better describe a temporary property of the effector of the grasp, usually a particular hand shape expressed by means of similes and analogies. Strikingly, the data presented in the previous paragraphs show that all these words denote instruments (in particular, containers) and are almost entirely found in the descriptions provided for substances and aggregates (cf. Table 4.7, p. 154). This is not by mere chance: as already discussed (cf. §4.6.2), the water, the flour and the sand are entities without a solid surface that can be grasped securely and held in a hand («ci infilo le mani sotto, ma visto che è liquida non riesco a fermarla»). Therefore, an instrument is required, in order to control such entities; in particular, a container.

The most interesting aspect is that, during the task, it was never stated that participants should refer only to a body part as the effector of the described grasp. Informants were only required to imagine a situation in which they had to grasp the entities they were seeing on the pc monitor, and to describe the action they would have performed. As a matter of fact, there were some students that simply said that they would have used a container. However, most participants variously described a grasp performed with the cupped hands: they somehow attributed to their hand the *properties* of an instrument.

The most evident property that a cupped hand shares with a real cup is at the perceptual level: the hand assumes a configuration such that its fingers and its palm together form a curved shape, which looks like that of a cup. But here, the most important property shared by the cup and the hand is probably rooted at a deeper level: both the effector and the object named by informants are the *instrument* through which the action of grasping may be realised. Therefore, it is not only the shape of the cup, but crucially its function (i.e. the function to contain something), the common ground on which the comparison between the hand and the cup can be established. This is the reason why, leaving aside *coppa* (attested nine times, also because of the collocation *mani a coppa*, "cupped hands"), the hand may be likened to a container, a spoon, or a ladle (*contenitore*, *cucchiaio*, *mestolo*, three occurrences each). It seems that it is the functional property of these objects, more than their perceptual properties, that allows the comparison with the hand. Otherwise, we would not find a mention of a simple instrument (*utensile*, two occurrences), e.g. «con tutta la mano, usandola come un mestolo insomma, un utensile». It is probably not by chance if in this example, as in other ones, the verb associated to the effector is *usare* or *utilizzare*, "to use" (e.g. «usando la mano a coppa, come un utensile»; «usando una mano come una coppa»).

In the grasp descriptions provided by informants, the hand is generally presented as the instrument through which the action is performed by the agent and mostly occurs in a *with*-phrase. The more the entity to be grasped gets far from affording an easy, secure, and stable manual grip, the more the hand acquires the properties of the instrument artefact most suited for that circumstance (here, a container). A language rich in similes and analogies reflects the overlapping between these two spheres, that of a "hand-effector" and that of a "container-effector". This process may also lead to a complete substitution of the body part-effector with an instrument artefact. Notably, this only happens for the flour and the water, the entities that afford the most difficult (if any) manual interaction.

It is worth noting that the similarity between the effector hand and the real container is reflected also at a syntactic level. Even if a detailed analysis of the syntactic structure of the transcripts has not been conducted, it is evident (even from the examples cited so far) that, in the vast majority of cases, the effector of the grasp occurs in a *with*-phrase and plays an Instrument role (e.g. «afferrerei il manico con una mano»; «il pallone lo prenderei con due mani»), exactly as the containers (e.g. «la farina la prenderei con un bicchiere»; «questa, mi aiuterei con una tazza, per raccoglierla»). Only in very few cases the effector of the grasp (always a body part, never a real container) occurs as the subject of the sentence. For example, consider this description for the box: «due mani di lato si avvicinano, stringono e sollevano».[45] In this specific linguistic representation of the event, the hands are not presented as instruments (for the notion of subject instrument, see for example Alexiadou and Schäfer 2006); rather, they are simply the participants involved in the event described by the sentence and are not modified by the event; therefore, they seem to fulfill the Theme role.

The Instrument semantic role presents differences in definition and causes problems of attribution, especially when an inanimate entity occurs as subject of a sentence. In this regard, Varvara and Ježek (2014: 387) highlight «the importance of distinguishing between semantic roles - relational notions belonging to the level of linguistic representation - and ontological types, which refer to internal qualities of real-world entities». Examples they mention are, for Italian, *la penna scrive nero* ("the pen writes black"), *forbici che tagliano bene* ("scissors that cut well"), in which "pen" and "scissors", typically occurring as Instrument in a *with*-phrase, are the subjects of the two sentences. The authors argue that inanimate nouns denoting instruments in subject position are not instantiations of the Instrument role, but of the Cause, Agent or Theme role. In this regard, also in «due mani di lato si avvicinano, stringono e sollevano», *mani* is presented as a Theme (cf. also Ježek et al. 2014; Ježek and Varvara 2015).

#### 4.7.1.4 Merging categories for effector-related words

The categories into which effector-related words have been classified might be gathered together into four macro-categories. The first one collects words that refer to a body part,

[45] Cf. descriptions such as the following ones (provided for the box and the pacifier respectively), in which the verbs *stringere* and *sollevare* are used: «la solleverei con entrambe le mani, lateralmente»; «dalla parte sotto, via, non quella che va in bocca, la stringerei con indice e pollice forse anche col medio».

and results from the fusion of the hand (hnd), meronym (mer), holonym (hol) and other body part (obp) categories (cf Tab. 4.7). The second one assembles spatial (spa) and perceptive (per) expressions into a single visuo-spatial class. The third one is that of quantity (unvaried, with respect to Tab. 4.7). Finally, the fourth macro-category groups the effector-related words denoting instruments (mostly containers), which in the vast majority of cases are simply evoked to describe the shape of the hand (spp), but sometimes are explicitly indicated as the instrument with which the grasp described is performed (ins).

| | | body part | visuo-spatial | quantity | instrument | tot. |
|---|---|---|---|---|---|---|
| Humans | freq. | 43 | - | 26 | - | 69 |
| | **%** | **62.3** | - | 37.7 | - | |
| Artefacts with AP | freq. | 655 | 119 | 204 | 2 | 980 |
| | **%** | **66.8** | 12.1 | 20.8 | 0.3 | |
| Artefacts without AP | freq. | 335 | 23 | 166 | - | 524 |
| | **%** | **63.9** | 4.4 | 31.7 | - | |
| Natural kinds | freq. | 141 | 9 | 52 | - | 202 |
| | **%** | **69.8** | 4.5 | 25.7 | - | |
| Substances/ Aggregates | freq. | 149 | 12 | 51 | 38 | 250 |
| | **%** | **59.6** | 4.8 | 20.4 | 15.2 | |
| Tot. | freq. | 1323 | 163 | 499 | 40 | 2025 |
| | **%** | **65.3** | 8.1 | 24.6 | 2 | |

Table 4.10. Merged classification of effector-related words with relation to the different categories of objects-stimuli: frequencies data and percentages (over the total number of effector-related words extracted for each category).

These frequency data can be also visualised in the mosaic chart[46] represented in Fig. 4.2, in which the five categories of objects-stimuli are identified by different colours and the area of each rectangle that composes the chart is proportional to the observed frequency in the corresponding cell (cf. Tab. 4.10; the values are reported in the graph's labels). A mosaic chart essentially combines a 100% stacked column chart and a 100% stacked bar chart in one single view. It works like a 100% stacked column chart (within each column, the height of the rectangles represents the proportion of the number of effector-related words observed for each stimulus-category over the total number of words pertaining to a specific domain);

[46] Graphs has been created with the Mekko Chart Creator add-in for Microsoft Excel (https://www.add-ins.com/).

additionally, the width of each column is proportional to the total value of the column with respect to the total number of words extracted.

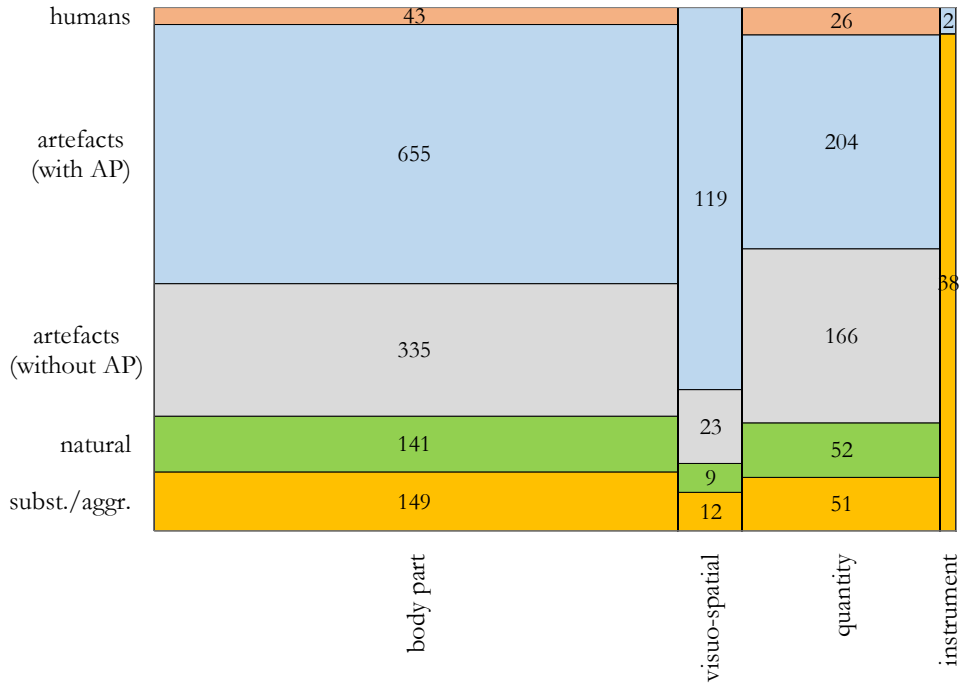



Figure 4.2. Mosaic chart representing the merged classification of effector-related words extracted from the descriptions provided for the different categories of stimuli. The labels report frequency values.

This mosaic well shows the similarities and the differences among the five categories of stimuli that we have already discussed in the last paragraphs. We observe that the mentions of body parts are the most frequent for all objects-stimuli (cf. §4.7.1.1), followed by modifiers expressing quantity. However, visuo-spatial expressions are more frequently produced for artefacts with affording parts than for other stimuli (they are never produced for the human kinds; cf. §4.7.1.2), whereas references to containers are almost only found for substances and aggregates (cf. §4.7.1.3), except in two cases, found for the overturned coffee cup.

### 4.7.2 Words related to the target of the grasp

We can now turn to the analysis conducted on target-related words, the results of which are reported in the following table (both frequencies and percentages are shown).

| | | ent | mer | spa | qua | per | spp | aen | tot. |
|---|---|---|---|---|---|---|---|---|---|
| Humans | freq. | 3 | 91 | 33 | - | - | - | 8 | 135 |
| | % | 2.2 | **67.5** | 24.4 | - | - | - | 5.9 | |
| Artefacts with AP | freq. | 75 | 477 | 417 | 8 | 21 | 17 | - | 1015 |
| | % | 7.4 | **47** | 41.1 | 0.8 | 2.1 | 1.6 | - | |
| Artefacts without AP | freq. | 37 | 12 | 148 | 8 | 9 | 1 | - | 215 |
| | % | 17.2 | 5.6 | **68.8** | 3.7 | 4.2 | 0.5 | - | |
| Natural kinds | freq. | 12 | 9 | 39 | 5 | 1 | 18 | - | 84 |
| | % | 14.3 | 10.7 | **46.4** | 6 | 1.2 | 21.4 | - | |
| Substances/ Aggregates | freq. | 15 | 3 | 14 | 34 | - | 5 | - | 71 |
| | % | 21.1 | 4.2 | 19.7 | **47.9** | - | 7.1 | - | |
| Tot. | freq. | 142 | 592 | 651 | 55 | 31 | 41 | 8 | 1520 |
| | % | 9.3 | 39 | **42.9** | 3.6 | 2 | 2.7 | 0.5 | |

Table 4.11. Classification of target-related words with relation to the different categories of objects-stimuli: frequencies data and percentages (over the total number of target-related words extracted for each category).

It is evident that, comparing the five categories of objects-stimuli, there are many considerations worth discussing.

#### 4.7.2.1 Meronyms and spatial relations

In the analysis conducted in the previous chapter, we considered as true indicators of the target of the grasp expressions that denote the part of the object with which the effector comes in contact (§3.1.1). Such parts may be directly referred to in two different ways: either by words pertaining to the spatial domain, which somehow refer to the object's subspace to which the effector is directed, or by words indicating objects' parts (meronyms). The results of the analysis reveal an important difference among the objects' categories (cf. §3.7): the more an object-stimulus is on the left part of the hierarchy *humans* > *artefacts with affording parts* > *artefacts without affording parts* > *natural kinds* > *substances and aggregates*, the more likely the target of the grasp is to be named in the descriptions.

The stimuli for which the target of the grasp is most frequently mentioned are, indeed, the human kinds and the artefacts with affording parts (cf. Fig. 3.10, p. 97).

Now, we are able to go a step further and notice that these two categories are also those in which meronyms are the most frequent type of word referred to the target of the grasp. In particular, for the human kinds, most participants described a grasp directed to a body part (cf. §4.4.2) and the number of meronyms exceeds that of spatial expressions, as it appears evident also from the mosaic chart reported in Fig. 4.3 (based on frequency data presented in Tab. 4.11). On the contrary, spatial expressions are much more frequent than meronyms for artefacts without affording parts, natural kinds, substances and aggregates:[47]

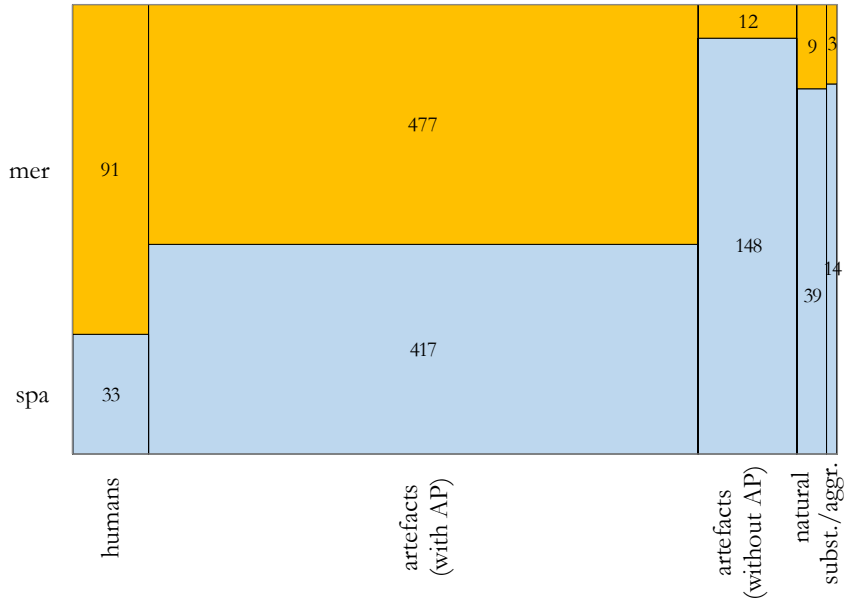



Figure 4.3. Mosaic chart representing the target-related words classified within the *space* and *meronym* categories extracted from the descriptions provided for the different groups of stimuli. The labels report frequency values.

For what regards artefacts with affording parts, presented with or without different orientation, we observe that even if the target of the grasp is in most cases indicated by lexemes that denote specific object's parts (cf. §4.3.3.2), spatial expressions are almost as frequent as meronyms.

A possible explanation for this is that all the parts of an object that can be denoted by meronyms, in theory, could also be indicated by visuo-spatial expressions, without compromising the clarity of the description provided. For instance, a grasp directed to the

[47] With respect to the previous chart presented in §4.7.1.4, in this mosaic chart (as well as in the next one, presented in §4.7.2.3) the *x*-axis and the *y*-axis were switched in order to enhance the readability of the graph.

*handle* of a rightward-oriented jug might be also indicated as a grasp directed to *the right part* or *the right side* of the jug, or more precisely to *the thinner part, on the right side* of the object. The contrary does not hold. Therefore, spatial expressions are sometimes used to indicate an object's part for which there is, in Italian, a specific name, but it does not come to the participant's mind during the interview. Similarly, participants might also prefer to use a spatial expression when they are not sure about how to name a specific object's part (because there is more than one possible word), or in order to avoid ambiguity (when an object has more than one part likely to be grasped).

When informants explicitly mention the target of a grasp directed to an artefact without affording parts or a natural kind, they usually resort to lexical expressions pertaining to the spatial domain, either because a specific object has no visually distinguishable parts, or because the parts denoted by the meronyms do not afford grasping (as for the stalk of the apple and the banana, or the handle of the overturned cups). This is the reason why words expressing spatial relations are the most frequent, among those found with reference to such stimuli. However, in these cases, participants are also less likely to explicitate a reference to the target: the descriptions they provide often lack this information and simply contain a reference to the effector (cf. §3.7).

#### 4.7.2.2 Substances and aggregates

In the analysis presented in Chapter III (in particular, cf. §3.5), we observed that the descriptions provided for substances and aggregates were characterised by a very low number of references to the target of the grasp. This is confirmed also by the present analysis: words that pertain to the spatial domain, as well as references to meronyms, are much rarer with comparison to other objects' categories. Interestingly, the types of words more frequently found with reference to substances and aggregates are expressions of quantity, followed by words denoting the entity itself (cf. Tab. 4.11, p. 164). As already discussed (cf. §4.2), these expressions do not indicate in a precise way the target of the grasp, but still their distribution is meaningful.

Expressions of quantity referred to the object-stimulus never occur for human kinds, and they are very rarely found in the descriptions provided for artefacts, both with and without affording parts (16 occurrences for a total of 31 objects-stimuli, i.e. 930 grasp descriptions). In the transcripts collected for the four natural kinds, quantity expressions related to the

target of the grasp are slightly more frequent and constitute the 6% over the total number of target-related words extracted for this category. However, for substances and aggregates they almost constitute half of the words with which participants referred to the flour, the water, the sand and the pumpkin seeds.

This distribution reminds us the hierarchy already discussed with relation to the probability for the target and the effector of the grasp to be named (*humans* > *artefacts with affording parts* > *artefacts without affording parts* > *natural kinds* > *substances and aggregates*). Interestingly, expressions denoting the entity itself (in most cases occurring as a direct object of a verb of grasp), as quantity expressions, are the least frequent in the descriptions provided for the human kinds, and the most frequent in those provided for substances and aggregates.

Probably, it is the difficulty in finding a way to describe the grasp of the water, the sand, the flour or the mound of pumpkin seeds what leads the informants to provide a sort of additional information in the descriptions. If they are not able to specify *where*, or towards *which part*, they would direct their hand during the grasp (since substances have no easily discernible and/or cognitively salient parts),[48] they can only simply repeat the information already provided by the experimental setting and name the object they are seeing (e.g. *I would grasp… the flour*), otherwise specify *how much* water, flour, sand, and *how many* seeds they would grasp.

#### 4.7.2.3 Merging categories for target-related words

For the target-related words, too, we can now merge the different semantic classes into four macro-categories.

The first one is the part category and collects meronyms (mer) and associated entities (aen), because concrete nouns pertaining to these two categories denote the exact part of the stimulus represented to which a grasp might be directed. A single visuo-spatial category collects words pertaining to the perceptive and spatial domains. The other classes are unvaried.

[48] Cf. §2.4.3.2 and, in particular, §4.6.1-2.

| | | part | visuo-spatial | quantity | entity | spp | tot. |
|---|---|---|---|---|---|---|---|
| Humans | freq. | 99 | 33 | - | 3 | - | 135 |
| | % | **73.4** | 24.4 | - | 2.2 | - | |
| Artefacts with AP | freq. | 477 | 438 | 8 | 75 | 17 | 1015 |
| | % | **47** | 43.2 | 0.7 | 7.4 | 1.7 | |
| Artefacts without AP | freq. | 12 | 157 | 8 | 37 | 1 | 215 |
| | % | 5.6 | **73** | 3.7 | 17.2 | 0.5 | |
| Natural kinds | freq. | 9 | 40 | 5 | 12 | 18 | 84 |
| | % | 10.7 | **47.6** | 6 | 14.3 | 21.4 | |
| Substances/ Aggregates | freq. | 3 | 14 | 34 | 15 | 5 | 71 |
| | % | 4.2 | 19.8 | **47.9** | 21.1 | 7 | |
| Tot. | freq. | 600 | 682 | 55 | 142 | 41 | 1520 |
| | % | 39.5 | **44.9** | 3.6 | 9.3 | 2.7 | |

Table 4.12. Merged classification of target-related words with relation to the different categories of objects-stimuli: frequencies data and percentages (over the total number of target-related words extracted for each category).

Again, we can present these data with a mosaic chart (Fig. 4.4):

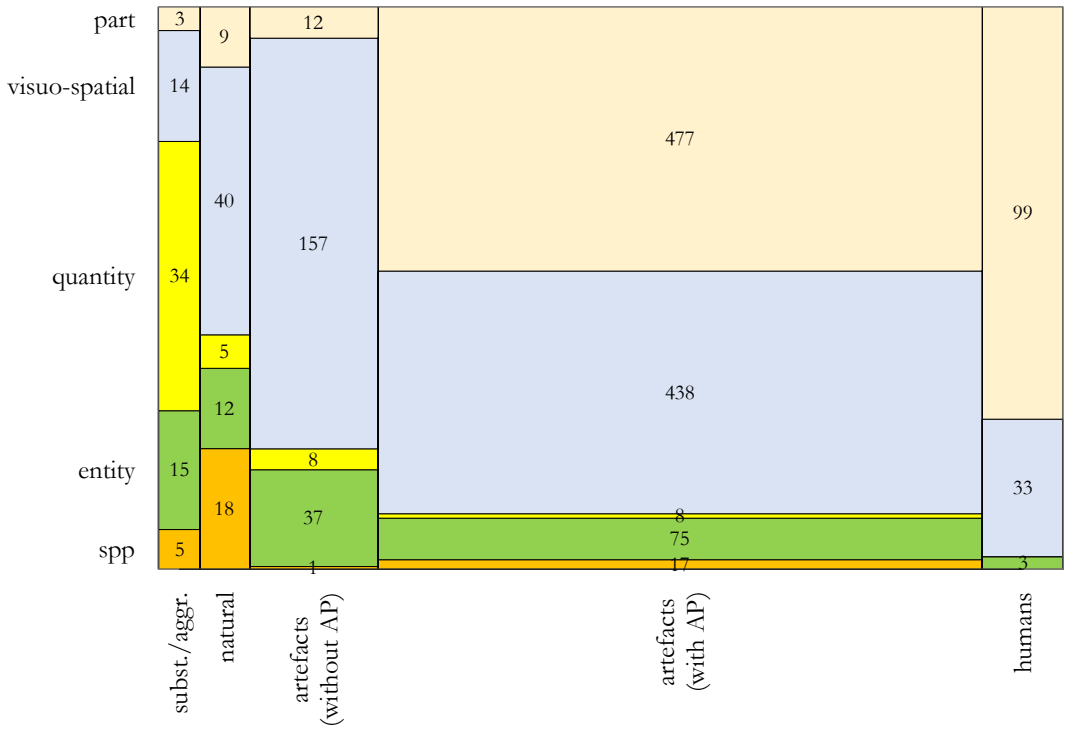



Figure 4.4. Mosaic chart representing the merged classification of target-related words extracted from the descriptions provided for the different categories of stimuli. The labels report frequency values.

Words pertaining to the visuo-spatial domain (light blue rectangles) are, in absolute, the most frequently named with relation to the target of the grasp, especially for artefacts without affording parts (73%) and natural kinds (47.6%). Parts are rarely mentioned for substances and aggregates, as observed in the last paragraph, as well as for natural kinds and artefacts without affording parts, whereas they constitute almost half of the target-related words extracted for artefacts with affording parts (47%) and the vast majority of those extracted for the human kinds stimuli (73.3%). The target of a natural entity, but especially of a substance or an aggregate, is indeed simply described with a more generic reference to the entity itself or to similar entities (the green and orange rectangles respectively) mostly pertaining to the same category of stimuli. Number words and quantifiers (represented by yellow rectangles) are most frequently found with relation to substances or aggregates (as discussed in §4.7.2.2).

## 4.8 Concluding remarks

A growing body of research conducted in the behavioural and neurophysiological fields demonstrates the close connection between an agent's ability to perceive an object and possibilities for action. In particular, the discovery of canonical neurons circuits, which fire both when an agent is performing an action upon an object, and when the agent simply looks at an object related to action, indicates that perceiving some properties of manipulable entities activates a sort of action simulation in brain circuits (e.g. Grèzes, Armony et al. 2003). For such motoric representations, it is important not only how familiar an agent is with the objects themselves, but also, notably, how familiar an agent is with the actions in which objects are typically involved (e.g. Buccino et al. 2009; Gentilucci 2002). It can be stated, following Gentilucci (2002: 1152), that «from a motor point of view, familiar objects can be represented by the type/types of interaction that we habitually have with them».

This is the reason why artefacts and tools are more effective in activating motor simulations of grasping and manipulation with respect, for example, to natural kinds. And this is also the reason why, within the artefacts category, there are specific object features, such as the presence of what we have called an affording part, or the orientation of such part, that are able to influence motor responses. A handle is the part of an object that humans usually grasp; therefore, for a sort of memory of past interactions with the object, the same

motoric patterns are activated both during the visualisation of a handled object, as well as during object grasping. We can recall, in this regard, the experiment conducted by Buccino et al. 2009, in which the largest motor evoked potential area was recorded from hand muscles when subjects looked at images of right-handled cups; if the handle of the object was located on the left side of the object or was broken, the motor activation was much less evident.

The results emerging from the experiments discussed in these chapters comply well with these findings. In particular, the analysis conducted in Chapter III reveals that the target of the grasp is named in most grasp descriptions provided for artefacts and human kinds, whereas it is less frequently mentioned for substances, aggregates and natural kinds. Moreover, within the category of artefacts, artefacts provided with affording parts are most frequently referred to as the target of the grasp, with respect to artefacts without affording parts. In light of the analysis conducted in the present chapter, we also notice that participants, especially for the category of artefacts with affording parts, usually refer to the target of the grasp by producing meronyms, i.e. by explicitly indicating and naming parts of the stimulus presented. In the frame of an action description task built around two components, i.e. the agent, who imagines performing an action, and the visually presented object, towards which that imaginary action is directed, the greater number of linguistic references to the target and a corresponding less frequent mention of the effector of the grasp, together with the explicit mentions of meronyms, are symptoms of a shift of attention, which from a focus on the sphere of the subject slides towards a focus on the sphere of the object. This shift, that is most evident for the two categories of artefacts, in comparison to natural kinds, turned out to be sensitive to the same factors that behavioural and neurophysiological research indicates as able to influence motor responses (namely the objects' semantic type, the presence of affording parts); therefore, also agents' linguistic behaviour does reflect affordances.

Interestingly, this study also takes into consideration two different kinds of stimuli that were not considered in the literature reviewed in Chapter I. Results obtained for the category of the human kinds are similar to those obtained for artefacts with affording parts. They usually elicit the most frequent references to the target of the grasp, with respect to references to the effector, this target being usually indicated by meronyms. Artefacts with affording parts, as well as humans, are the kinds of entities with which humans most frequently interact, and for which the effect of past interactions is probably stronger. We generally take jugs and cups by their handles, and humans by hands or by arms, because these actions correspond

to what we habitually do with such entities, which also are very meaningful for us: most of our everyday manual actions (and in particular grasping) directly involve or are directed to other humans or to artefacts. On the contrary, our manual interaction with natural kinds, and in particular with their parts, is much less frequent. We rarely take mandarins and bananas by their stalks, and we usually do not need to. The importance of the kind of interactions we usually have with the environment in which we live, as well as with the entities that belong to it, also seems to explain results from substances and aggregates, for which the target of the grasp is less frequently named.

Therefore, it is not only the mere presence *versus* the absence of a clearly distinguishable part of an object-stimulus that explains the results of the present analysis: this argument could only explain why *parts* are frequently named as the target of the grasp for artefacts and humans, especially with meronyms, but it does not fully explain why there are a few mentions of the target of the grasp for other kinds of stimuli (especially natural kinds and substances). In theory, *every description* could contain a reference to the target of the grasp, expressed at least by referring to spatial relations.

There is at least another factor, strictly connected to objects' familiarity and habitual interactions, which could be taken into account, i.e. that the target is more likely to be named, either by meronyms or by spatial expressions, when the stimulus presents *different* and *competing* possible targets associated to different habitual, highly familiar actions. A part or a subspace of an object is more likely to be named especially if there is another part or another subspace of the object (usually more than one) that could constitute a target for a *meaningfully different* kind of grasp, i.e. a grasp that is preliminary to performing a different subsequent habitual action and to reach a different goal. For instance, the target of the grasp is rarely mentioned for the stone. The stone has no distinguishable and protruding part, it has a continuous contour and a regular surface; the right side of a stone does not afford a less effort-consuming, more stable, or safer grasp, with comparison to its upper part, or to its left part. Moreover, and most importantly, the different kinds of grasps that could be performed on a stone are not usually associated to different kinds of actions, and this may be another reason why they are rarely indicated in linguistic descriptions by means of spatial terms. On the contrary, for artefacts with affording parts and humans, it seems that not only these stimuli have visually distinguishable parts that could constitute a good (e.g. less effort-consuming, more stable, safer) target of the grasp, but also that the grasp of one or another specific part is preliminary to different kinds of actions. For the overturned coffee cup, as

well as for the two overturned tea-cups, the handle is rather rarely mentioned, with comparison to other artefacts with affording parts. On the one hand, this confirms that is not the presence of a visually perceivable part, *per se*, what attracts the grasp (cf. the rare mentions of the apple's and the banana's stalks). Nevertheless, most informants named the target (usually the upper part of the overturned cup), for these three stimuli.

This probably happens because participants recognise *more possible targets* of the grasp (in particular, the grasp offered by the handle), each one affording different actions: for instance, a grasp from the above part would be more suited for moving the overturned cup, a grasp from the handle would permit to turn the object upside down. Interestingly, this argument is in line with the hypothesis that the motor representation evoked by a single object codes *all* affordances enabled by the stimulus. The fact that different possibilities for actions (in this case, more than one grasp type) are simultaneously activated by a visually presented object is supported by empirical results. For instance, we have already discussed in §1.3.1 an experiment conducted by Gentilucci (2002: 1150-1151), in which subjects grasp two bells of different volume in different ways, even when their stalks, that is the part where the participants are explicitly requested to grasp the two objects, have an identical dimension. The hand, when the target-object is the large bell, is partly preshaped as if the entire object, and not only its stalk, was to be grasped.

# Chapter V

# Affording Parts and Objects Representations

The analysis illustrated in the previous chapters has focussed on the reflexes of affordances in language. In particular, the experiment conducted has shown that the target of the grasp is more likely to be named with respect to the effector of the grasp when the object-stimulus presented to the subject is an artefact or a human kind. Artefacts constitute a category of particular interest, allowing us to compare the results from the present analysis with the research conducted on grasping and manipulation within the neurophysiological, neuropsychological and behavioural fields (for which cf. Chapter I). We observed that affordances effects on linguistic production are more evident when the stimulus presents affording parts: therefore, linguistic behaviour is likely to reflect a focus of attention on the object stimulus; in particular, on the object's part that attracts the grasp. Moreover, the analysis presented in §4.7.2.1 has shown that the target of the grasp of artefacts with affording parts is in most cases named with meronyms.

However, we do not know whether the frequent reference to objects' meronyms, which emerged from the analysis illustrated in the last chapter, is mostly due to the specificity of

the task (that was strictly focussed on the act of grasping), or may be also due to the fact that objects' meronyms (and, in particular, some specific meronyms) constitute an important semantic feature for the concept of the object-stimulus *per se*.

In order to address this issue, a second study was conducted on the 14 stimuli of artefacts with affording parts. A very different methodology has been adopted, based on a property generation task (semantic feature production). The following paragraphs will introduce semantic norms (§5.1) and describe the feature production task conducted on the artefacts with affording parts stimuli (§5.2); then, a brief overview of the results of the study will be provided (§5.3); finally, the production of meronyms in the property generation task will be illustrated and compared to the production of meronyms in the grasp description task (§5.4, §5.5). In the last section (§5.6), the results of this comparison will be discussed.

## 5.1 Semantic feature norms: a brief introduction

In psychological and cognitive studies, semantic norms are collections of features produced by subjects in a property generation (or feature listing) task. Informants are typically presented with a written word (e.g. *airplane*), and they are asked to write a number of properties (e.g. *flies*; *has wings*) that they judge relevant to describe its meaning (Kremer and Baroni 2011). Once data have been acquired, they undergo a process of normalization[49] and the semantic features produced are finally classified according to a given set of semantic types.[50] After this process of normalization and classification, a distributional analysis may be conducted on feature norms in order to highlight, for example, which kinds of (and how many) features have been produced for a given stimulus, or how many informants produced that particular feature; eventually, additional statistical analysis may be conducted (McRae et al. 2005).

[49] Normalization is necessary because there is a great variability in the way subjects may list features; cf. §5.2.3.

[50] For instance, the properties *takes off*, *lands*, *flies*, referred to an airplane, are related to the events in which the word's referent is typically involved, whereas *has wings*, *has a motor* are related to the parts of which it is constituted; cf. §5.2.4.

Researchers typically assume that property generation tasks open a window on the representation of a concept that underlies word meaning[51] (e.g. Wu and Barsalou 2009: 174; Kremer and Baroni 2011: 98). According to the embodied cognition view, concepts are deeply rooted in sensory-motor activity and are retrieved through the re-enactment of the concrete experiences people had with real entites (Gallese and Lakoff 2005; Barsalou 1999; 2008; Borghi 2007). This can also be true when concepts retrieval is mediated by words (Borghi 2007).[52] Following from this account, the features listed in a property generation task are generally considered as deriving from the concrete interactions of the subject with the referent of the word-stimulus. For instance, Cree and McRae (2003: 167) assume that «when participants call to mind features to list in the norming task, they directly tap into representations that have developed through repeated multisensory exposure to, and interactions with, the various objects».

However, the nature of semantic norms is not totally clear, also because cognitive and neural mechanisms underlying feature listing have rarely been directly investigated (Santos et al. 2011: 84). Moreover, as already stated, the most typical way to derive semantic norms is through presenting subjects with a written word and asking them to describe its meaning. Therefore, informants access the concept in a very specific way, i.e. through a particular lexeme. Interestingly, Santos et al. (2011) argue that feature production originates in two distinct (although interacting) systems, i.e. the linguistic form system and the situated simulation system. According to the authors, the simple perception of a word is able to elicit other associated linguistic elements. For instance, the presentation of the word *car* might elicit associated words such as *automobile* and *vehicle*. Then, the word's presentation also activates correlated simulations in the brain's modal systems:

[51] Nevertheless, these two levels, that of the lexical meaning, on the one hand, and that of the concepts, on the other, have to be set apart from one another: as Murphy notes, «there is a mismatch between words and concepts» (Murphy 2010: 38-39).

[52] We have already discussed (cf. §1.4.1-2) recent studies that suggest that language, too, is grounded in perception and action: the comprehension and the processing of action-related nouns, and in particular of words denoting graspable tools or artefacts, involve the activation of the motor system (e.g. Cattaneo et al. 2010; Gough et al. 2012; Marino et al. 2013). Such modulation of the cortical motor regions during noun words presentation is comparable to the activity observed during the visual perception of the corresponding objects or their images (in particular, cf. Shinkareva et al. 2011).

> We assume that the simulation system becomes active very quickly once the presented word form is recognized, but that the activation of a simulation proceeds more slowly than the activation of associated linguistic forms. By "simulation" we mean that the brain simulates the perceptual, motor, and introspective states active during interactions with the word's referents […] Recognizing the word "cat," for example, reenacts neural states that represent how cats look, sound, and feel, how one interacts with cats, and how one feels affectively (Santos et al. 2011: 88).

Therefore, in a feature listing task, both linguistic associations and simulations are probably involved (without excluding the contribution of other possible strategies): early properties have a linguistic relation to the word-stimulus and tend to originate in a word association process, whereas properties produced later describe, for instance, associated objects and situations, and tend to originate from situated simulations. Evidence in favour of this account comes from the results that the authors report from two experiments, in which responses linguistically associated to the word-stimulus (originating in the linguistic system) are produced earlier, whereas object-situation responses (originating in the simulation system) tend to occur later (cf. also Barsalou et al. 2008).[53]

There are some limitations about collecting and using semantic norms that are widely discussed in the literature (McRae et al. 2005; Kremer and Baroni 2011; Lenci et al. 2013). First, since the features are linguistically produced (either in a written or verbal form), it could be the case that a particular aspect of a meaning is underrepresented in the norms, not because it is not salient for a given concept, but simply because it is not easy to express verbally. Second, it has been observed that subjects tend to produce more features that are effective in distinguishing a concept from the others of the same category, compared to

[53] The fact that linguistic and object-situations responses originated in two different systems is also confirmed by a companion neuroimaging study (Simmons et al. 2008). In particular, the study revealed that, during a property generation task, the areas more active for word association than for situated simulation included the left inferior frontal gyrus (Broca's area), the left temporal gyrus, and the right cerebellum, i.e. regions that are involved in word processing, especially word production. On the contrary, the areas more active for situated simulation included the precuneus, the right middle temporal gyrus, and the right middle frontal gyrus, i.e. regions often associated with various forms of simulation that underlie mental imagery, episodic memory, and situational processing (Santos et al. 2011: 110-111).

features that are true for a large number of concepts.[54] Nevertheless, in the last decades semantic norms collected with a property generation task have been widely adopted in cognitive sciences: for instance, they can be used to test theories of semantic memory, to develop computational models, as well as for a multitude of other purposes (see, for instance, McRae et al. 2005: 548; Kremer and Baroni 2011: 98; Lenci et al. 2013: 1220).

Semantic norms have been collected for many languages, for concrete or abstract nouns, as well as for verbs (e.g. McRae et al. 2005; De Deyne et al. 2008; Vinson and Vigliocco 2008; Kremer and Baroni 2011; Frassinelli and Lenci 2012; Montefinese et al 2013; Lenci et al. 2013). These works differ (to a lesser or greater extent) from one another in the procedure adopted to collect, normalise, and classify features: the methodology adopted in the present study that is described in the following paragraphs is inspired by McRae et al. (2005), Kremer and Baroni (2011) and, in particular, by Lenci et al. (2013).

## 5.2 Methods

### 5.2.1 Participants and stimuli

Thirty young participants entered the study, 24 females and 6 males. They were all native Italian speakers, and none of them had participated in the action description task. For the most part, they were undergraduate students enrolled in the University of Pisa and the University of Genoa; only four of them had already graduated. Their age was between 21 and 39 years (mean=27.2; SD=3.59).

Since this experiment has been expressly conducted to investigate whether the objects' meronyms linguistically produced in the previous test constitute an important semantic feature for the concept of an object-stimulus, lexical expressions denoting the 14 artefacts with affording parts (the category for which meronyms have been more frequently produced) were used as stimuli. Such lexical expressions were the following (the order of appearance is respected): *brocca* ("jug"); *tazza* ("cup"); *spada* ("sword"); *phon* ("hairdryer"); *microfono*

[54] However, this may be also a strong point: «general features play only a small role in object identification, language comprehension, and language production precisely because they are not salient and are true of large numbers of concepts» (McRae et al. 2005: 549).

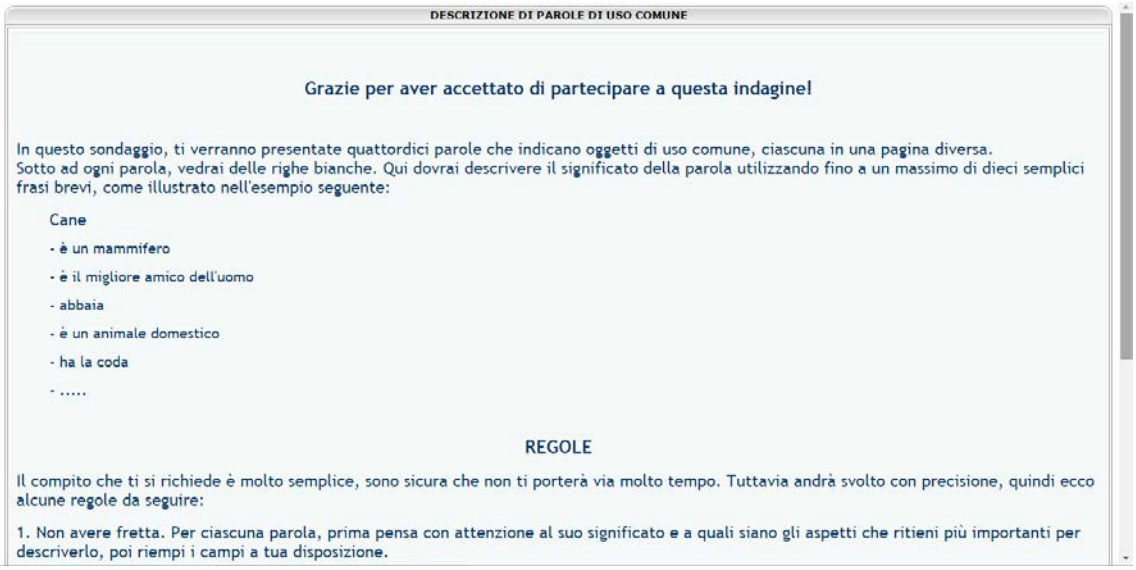
("microphone"); *borsa da donna* ("girl bag"); *ciuccio* ("pacifier"); *sedia* ("chair"); *trolley*; *canotto* ("rubber boat"); *zaino* ("backpack"); *ombrello* ("umbrella"); *mestolo* ("ladle").

### 5.2.2 The survey

The experiment was conducted on the Internet through a web interface created with the LimeSurvey software, a free open source survey tool made available by the Computational Linguistics Laboratory (CoLing Lab) of the University of Pisa.

Through an invitation email, in which the scope of the study was briefly introduced, participants could directly access the online questionnaire that was entitled *Descrizione di parole di uso comune* ("description of common use words").

In the welcome page (a part of which is represented in Fig. 5.1), they could read the instructions on how to complete the survey, together with a brief example of semantic features listing.


DESCRIZIONE DI PAROLE DI USO COMUNE

Grazie per aver accettato di partecipare a questa indagine!

In questo sondaggio, ti verranno presentate quattordici parole che indicano oggetti di uso comune, ciascuna in una pagina diversa.
Sotto ad ogni parola, vedrai delle righe bianche. Qui dovrai descrivere il significato della parola utilizzando fino a un massimo di dieci semplici frasi brevi, come illustrato nell'esempio seguente:

Cane
- è un mammifero
- è il migliore amico dell'uomo
- abbaia
- è un animale domestico
- ha la coda
- .....

REGOLE

Il compito che ti si richiede è molto semplice, sono sicura che non ti porterà via molto tempo. Tuttavia andrà svolto con precisione, quindi ecco alcune regole da seguire:

1. Non avere fretta. Per ciascuna parola, prima pensa con attenzione al suo significato e a quali siano gli aspetti che ritieni più importanti per descriverlo, poi riempi i campi a tua disposizione.

Figure 5.1. Welcome page of the survey *Descrizione di parole di uso comune*.

The text reported in the welcome page[55] was divided into two sections. The first one presented the survey:

[55] An English translation of the instructions is provided in Appendix B.

Grazie per aver accettato di partecipare a questa indagine!
In questo sondaggio, ti verranno presentate quattordici parole che indicano oggetti di uso comune, ciascuna in una pagina diversa. Sotto ad ogni parola, vedrai delle righe bianche. Qui dovrai descrivere il significato della parola utilizzando fino a un massimo di dieci semplici frasi brevi, come illustrato nell'esempio seguente:
Cane
- è un mammifero
- è il migliore amico dell'uomo
- abbaia
- è un animale domestico
- ha la coda
- …

Since the survey was focussed on the artefacts category, the example provided was a stimulus of a different semantic class, i.e. an animal. The second part of the welcome page contained a few rules that were intended both to help the participant to complete the task correctly and to minimise the presence of free associations:

REGOLE
Il compito che ti si richiede è molto semplice, sono sicura che non ti porterà via molto tempo. Tuttavia andrà svolto con precisione, quindi ecco alcune regole da seguire:
1. Non avere fretta. Per ciascuna parola, prima pensa con attenzione al suo significato e a quali siano gli aspetti che ritieni più importanti per descriverlo, poi riempi i campi a tua disposizione.
2. Descrivi il significato della parola con frasi brevi. Cerca di essere chiaro e sintetico.
3. Non esistono risposte giuste o sbagliate: sei assolutamente libero di spiegare nella maniera che preferisci quello che ritieni essere il significato di queste parole.
4. Non sei obbligato a riempire tutte le dieci righe per ogni parola.
5. Una volta completata una pagina, accertati della correttezza dei dati che hai inserito, poiché non sarà possibile modificarli in seguito. Solo quando sei sicuro, clicca sul pulsante "Avanti", che ti porterà alla pagina successiva.
6. Non è consentito interrompere il questionario e salvare le proprie risposte: il questionario può essere salvato solo alla fine, dopo che è stato completato.

After this initial page, the following 14 pages contained only the title of the survey, the progress bar (which showed the percentage of the survey already completed), the word stimulus, and a short reminder of the task (see for example Fig. 5.2 and Fig. 5.3):

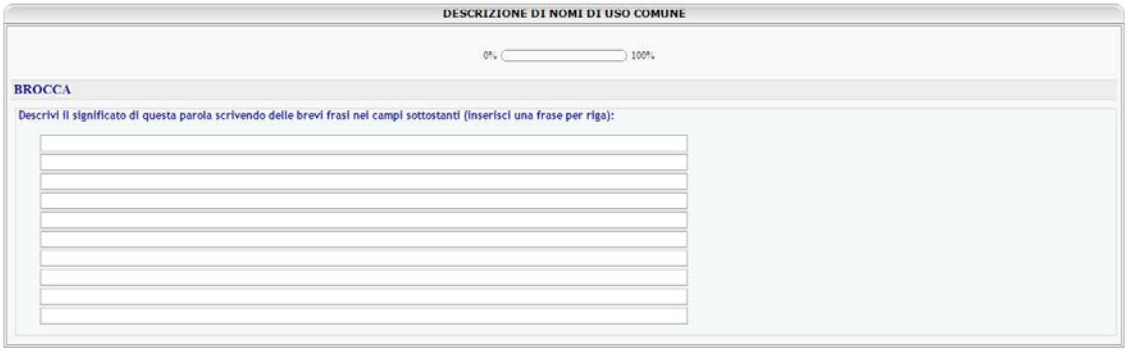



Figure 5.2. Example of a page of the survey *Descrizione di parole di uso comune*, for the word *brocca*, "jug".

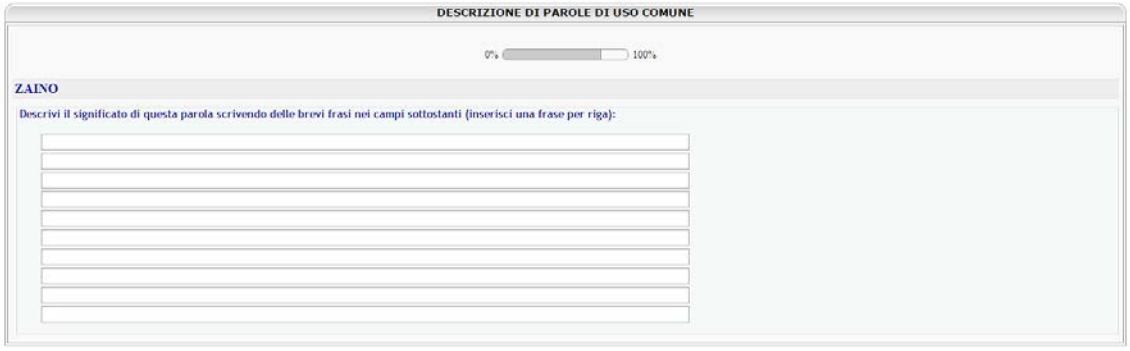



Figure 5.3. Example of a page of the survey *Descrizione di parole di uso comune*, for the word *zaino*, ("backpack").

Once the survey was completed, the participants received a confirmation email.

### 5.2.3 Normalization

All answers collected with the survey were imported in a single file and they underwent a long process of normalization, mostly inspired by Baroni et al. (2013) and Lenci et al. (2013) (for which cf. also Vinson and Vigliocco 2008; Kremer and Baroni 2011). This operation is necessary because there is an evident and wide variability in the descriptions produced, both within the group of participants and within the features produced by a single informant.

In particular, informants often filled the blank lines with long descriptions of stimuli that actually contained more than one property. In such cases, the descriptions were manually split into two or more single features, each one representing a separated nucleus of information. Table 5.1 provides some examples, taken from the descriptions collected for the word *brocca* ("jug"). A sentence like *è di vetro, terracotta, metallo o plastica* ("it is made of glass, earthenware, metal, or plastic"), actually contains four different pieces of information regarding a jug; therefore, each one was isolated and inserted in a separate line (lines 1-4). Usually, in this process of chunking, adjectives and nouns were split, as well as also verbs and their direct objects (lines 5-6, 7-8). There were even cases in which a single description was syntactically more complex and contained two or more clauses (lines 9-11). In this process of segmentation and simplification, synonymous features were merged into a single featural representation (compare lines 2 and 13).

Each feature was finally transformed into a normalised form (cf. Lenci et al. 2013: 1222). In this phase, nouns and adjectives were converted in their singular and masculine forms, and verbs in their active or passive infinitival forms (Table 5.1, column 4).

| Line | Description | Feature | Normalised Feature |
|---|---|---|---|
| 1 | È di vetro, terracotta, metallo o plastica | È di vetro | Vetro |
| 2 | È di vetro, terracotta, metallo o plastica | È di terracotta | Terracotta |
| 3 | È di vetro, terracotta, metallo o plastica | È di metallo | Metallo |
| 4 | È di vetro, terracotta, metallo o plastica | È di plastica | Plastica |
| 5 | Ha un collo largo | Ha un collo | Collo |
| 6 | Ha un collo largo | (Ha un collo) largo | Largo |
| 7 | Si usa per versare liquidi | Si usa per versare | Versare |
| 8 | Si usa per versare liquidi | (Si usa per versare) liquidi | Liquido |
| 9 | È un contenitore di vetro che si usa per versare | È un contenitore | Contenitore |
| 10 | È un contenitore di vetro che si usa per versare | È di vetro | Vetro |
| 11 | È un contenitore di vetro che si usa per versare | Si usa per versare | Versare |
| 12 | Può essere di vetro o di coccio | È di vetro | Vetro |
| 13 | Può essere di vetro o di coccio | È di terracotta | Terracotta |

Table 5.1. Example of feature normalization for five descriptions provided for the word *brocca* ("jug").

### 5.2.4 Classification

All normalised features were labelled with a semantic type according to the relation that each feature established with the stimulus for which it had been produced. For instance, *vetro*

("glass"), derived from the feature *è di vetro* ("is made of glass"), expresses the material of which a jug may be made, whereas *contenitore* ("container"), derived from the feature *è un contenitore* ("is a container"), expresses the category of objects to which a jug belongs.

There are many semantic annotation schemes proposed to classify features (compare, for instance, Cree and McRae 2003; McRae et al. 2005; Vinson and Vigliocco 2008; Wu and Barsalou 2009; Lebani and Pianta 2010; Kremer and Baroni 2011; Montefinese et al. 2013). The classification adopted in this study follows the annotation scheme described in Lenci et al. (2013), which is inspired by Wu and Barsalou (2009) and Lebani and Pianta (2010), but which considers a smaller number of feature types. Such scheme is articulated into five macro-categories (Lenci et al. 2013: 1222 ff.):

1) **taxonomical features**
2) **entity features**
3) **situational features**
4) **quantity features**
5) **introspective features**

In what follows, the feature types subsumed by each one of these macro-categories are reported, together with some examples of the stimulus-feature pairs.

The **taxonomic features** category includes superordinates (e.g. *brocca*, "jug" – *contenitore* "container"), coordinates (e.g. *mestolo*, "ladle" – *spatola*, "spatula"), subordinates and examples (e.g. *sedia*, "chair" – *sedia a dondolo*, "rocking chair"), and approximate synonyms (e.g. *brocca*, "jug" – *caraffa*, "pitcher").[56]

The second macroclass gathers **entity features**. It includes part-of relations, such as meronyms (e.g. *brocca*, "jug" – *manico*, "handle"), holonyms (e.g. *microfono*, "microphone" – *cellulare*, "mobile phone") and materials (e.g. *spada*, "sword" – *ferro*, "iron"); but also features referred to qualities that can be directly perceived (e.g. *spada*, "sword" – *lungo*, "long") or indirectly perceived (i.e. that cannot be apprehended with only direct perception and are often abstract; e.g. *borsa*, "bag" – *costoso*, "expensive").

[56] In Lenci et al. (2013) also antonyms (e.g. *freedom-slavery*) and instances (expressed by proper nouns, such as *mountain-Alpi*) were listed among the taxonomical features, but no relation of this type was found in the norms collected.

The third category is that of **situation features**, which contains features referred to the situations and contexts in which the stimulus is typically found, such as events (*tazza*, "cup" – *bere*, "drink"), concrete or abstract entities that are found in the same situation in which the stimulus is found (*mestolo*, "ladle" – *zuppa*, "soup"; *spada*, "sword" – *dolore*, "pain"), manners of performing an action associated with the stimulus (*tazzina da caffè* "coffee cup" – (*si usa*) *facilmente*, "(is used) easily"), locations (*canotto*, "rubber boat" – *mare*, "sea"), and temporal spans (*ombrello*, "umbrella" – *inverno*, "winter").

**Introspective features** express subjective evaluations (e.g. *sedia*, "chair" – *comodo*, "comfortable"), or emotions and feelings.

Finally, **quantity features** express a quantity (or amount) related to the stimulus or, more frequently, to one of its properties (e.g. *sedia*, "chair" – *quattro* (*gambe*), "four (legs)").

## 5.3 Semantic features produced

For the 14 stimuli considered, the 30 participants produced a total number of 2145 answers (mean= 153.2 answers per stimulus, SD= 12.26), which corresponds to a mean of 5.1 descriptions provided by each subject for one stimulus. All participants comprehended the task; however, within the total number of descriptions, 25 (1.2%) were excluded because they did not refer to an object's property. For instance, the informants sometimes provided a comment upon the morphological form of the word used as stimulus (e.g. *ombrello*: *è una parola derivata con suffisso*, "is a suffixed word"; *trolley*: *è una parola inglese*, "is an English word") or referred to homophones and polysemes (e.g. *phon*: *il suo nome é anche quello di un vento*, "it is also the name of a wind"; *spada*: *é anche il nome di un pesce*, "it is also the name of a fish"; *è anche uno dei quattro semi nel gioco di carte della briscola* "it is also one of the four suits in the briscola card game"). These are properties of the word-stimulus rather than properties of the concept for the entity it denotes. Such examples are of particular interest, since they support the idea that lexical processing plays an important role in feature listing tasks based on word presentation (cf. §5.1). Moreover, also idiomatical expressions were excluded (e.g. *spada*: *chi di spada ferisce, di spada perisce*, "he who lives by the sword, dies by the sword").

Considering the remaining 2120 descriptions, the total number of features produced is 3790 (844 distinct normalised features).

Each participant produced a mean of 126.3 features (there was a high variety in the number of features produced, ranging from 61 to 261 features, SD=49.37), which

corresponds to a mean of nine features per informant per stimulus. A mean of 270.7 features (SD=33.88) was produced for each lemma.

As already stated, this norming study is primarily aimed at establishing a comparison (for which cf. *infra*, §5.4) between the target-related words extracted from the grasp descriptions (i.e. the lexical expressions used to indicate the target of the grasp in the action description task presented in the previous chapter) and the semantic features produced by informants in the property generation task. This comparison will regard only 14 artefacts with affording parts (*brocca*, *tazza*, *spada*, *phon*, *microfono*, *borsa da donna*, *ciuccio*, *sedia*, *trolley*, *canotto*, *zaino*, *ombrello*, *mestolo*) and, in particular, will be focussed on the category of meronyms. Therefore, for the purpose of this thesis, only a very brief overview of the general results will be provided, leaving a more fine-grained analysis for future research.

For the category of stimuli here considered, we observe a strong predominance of features expressing situational properties (tot. 1918 features, 50.7%), followed by entity features (1370, 36.1%). The remaining features mostly express taxonomical relations (346, 9.1%), whereas properties related to quantity (81, 2.1%) or to subjective evaluations (75, 2%) are rather rare.

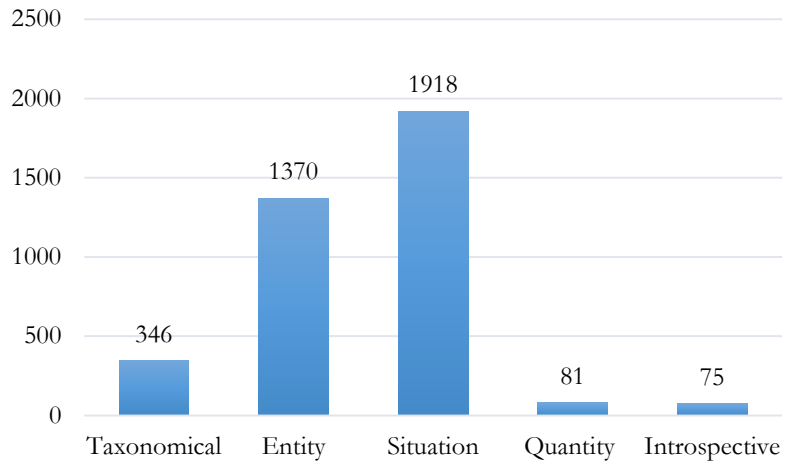



Figure 5.4. Results of the feature norms classification (tot. 3790 features) divided into five semantic macro-categories.

However, looking in more detail at the different kinds of relations within each macro-category, we can add further remarks.

The following table reports the general results of the features' classification (the feature types and the codes are the same as those adopted in Lenci et al. 2013), with the frequency of each feature type and its percentage over the total number of feature norms:

| Feature class | Feature type | Code | Freq. | % | Example |
|---|---|---|---|---|---|
| Taxonomic | Hypernym | isa | 284 | 7.5 | *brocca – contenitore* |
| | Example_of | exa | 30 | 0.8 | *sedia – sedia a dondolo* |
| | Coordinate | coo | 26 | 0.7 | *mestolo – spatola* |
| | Synonym | syn | 6 | 0.1 | *brocca – caraffa* |
| Entity | Meronym | mer | 327 | 8.6 | *brocca – manico* |
| | Holonym | hol | 10 | 0.3 | *microfono – cellulare* |
| | Made_of | mad | 395 | 10.4 | *spada – ferro* |
| | Perceptual property | ppe | 525 | 13.9 | *spada – lungo* |
| | Non-dicrectly perceptual property | pnp | 113 | 3 | *borsa da donna – costoso* |
| Situation | Event | eve | 885 | 23.4 | *tazza – bere* |
| | Entity (concrete) | eco | 652 | 17.2 | *mestolo – zuppa* |
| | Entity (abstract) | eab | 15 | 0.4 | *spada – dolore* |
| | Manner | man | 19 | 0.5 | *tazzina – (si usa) facilmente* |
| | Space | spa | 279 | 7.4 | *canotto – mare* |
| | Time | tim | 68 | 1.8 | *ombrello – inverno* |
| Quantity | Quantity | qua | 81 | 2.1 | *sedia – quattro (gambe)* |
| Introspective | Subjective evaluation | eva | 75 | 1.9 | *sedia – comodo* |

Table 5.2. Results of the feature classification (17 feature types): frequency and percentages over the total number of 3790 features extracted.

The distribution of feature types shown in Table 5.2 can also be represented in a bar graph (Fig. 5.5). The two categories of quantity and subjective evaluations are the least represented and collect, respectively, only the 2.1% and 1.9% of the features extracted. Among taxonomic features, the most numerous are hypernyms (7.5%); only 62 features (1.6%) establish with the word-stimulus a different kind of taxonomic relation. Perceptual properties (13.9%) and materials (10.4%) occupy the greatest part of entity features, followed by meronyms (8.6%).

The most numerous semantic class, however, is constituted by situational features, in particular events (23.4%, for which cf. *infra*) and, secondly, associated concrete entities (17.2%),[57] whereas spatial features occur more rarely (7.4%).

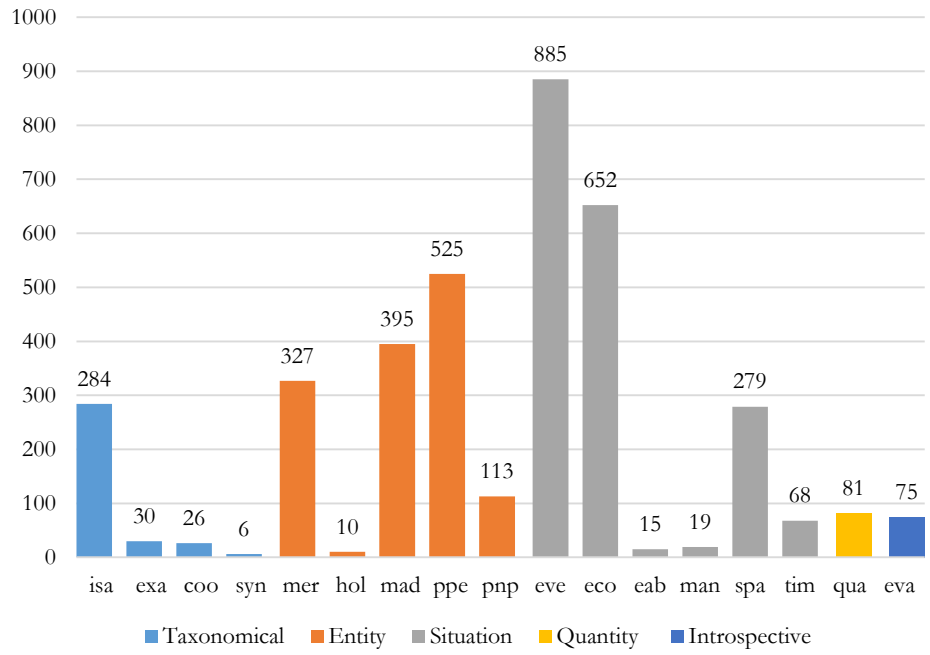



Figure 5.5. Results of the feature norms classification (tot. 3790 features) divided into 17 semantic categories (grouped into five macro-categories).

The category most represented for the artefacts stimuli, i.e. that of events, is worth a brief note. The highest frequency values scored by this class of features indicates that the most frequently named properties for the artefacts here considered are referred either to what an object is used for (i.e. its function), or to what an agent usually does with it. This is not surprising: artefacts can be defined as «physical objects that have been designed and made by human beings and that have both a function and a use plan» (Vermaas et al. 2011: 14): both objects' function and use-plan (a series of goal-directed actions to be met by the user to ensure that the artefact's function is realised) are typically represented by events and linguistically expressed by verbs.

[57] According to Wu and Barsalou (2009: 184), the fact that participants spent more time describing background situations than describing target objects supports the idea that they used situated simulations to generate properties.

Within the event category, considering the features with a frequency value greater than 15, at the top of the ranking we find *contenere*, "to contain", with 86 occurrences (17 occurrences for the jug, 15 for the girl bag and the cup, 13 for the trolley, 11 for the coffee cup and the backpack, four for the rubber boat).

Then, much lower in the ranking, we find the verbs *asciugare*, "to dry [sth]" (27 occurrences, for the hairdryer alone); *riparare*, "to protect" (26 occurrences, for the umbrella alone); *sedersi*, "to sit" (25 occurrences: 24 for the chair and only one for the rubber boat); *bere*, "to drink" (23 occurrences: 12 for the cup, 11 for the coffee cup); *trasportare*, "to carry" (20 occurrences: nine for the backpack, six for the trolley, three for the rubber boat, one for the jug and the ladle); *amplificare*, "to amplify" (18 occurrences, for the microphone alone); *emettere*, "to emit" (18 occurrences, for the hairdryer alone, referring to hot or cold hair); *essere_portato*, "to be carried" (18 occurrences: 11 for the backpack, three for the trolley and the umbrella, one for the girl bag); *essere_gonfiato*, "to be inflated" (18 occurrences, for the rubber boat alone); and finally *mescolare*, "to stir" (18 occurrences, for the ladle alone).

It is evident that, in most cases, these verbs express the primary function of the artefacts: typically, containers such as jugs and cups are used to contain liquids, hairdryers are used to dry hair, chairs are used to sit, ladles are used to stir, cups are used to drink, etc. Otherwise, features related to the use-plan are produced: for instance, a rubber boat needs to be inflated, in order to realise its function properly.

This kind of information cannot be compared to the results obtained from the action description task, since that experiment was strictly focussed on the action of grasping: subjects were asked to describe how they would have grasped the objects, and not to mention the actions that they would have performed with them.

We can now come back to the main topic of this analysis: presenting a detailed comparison between the target-related words extracted from the grasp descriptions commented in the previous chapter and the semantic features produced by informants for what regards the meronyms.

## 5.4 Meronymic features

For each word stimulus, a comparative table will report all meronyms produced in the property generation task, together with the target-related meronyms extracted from the grasp descriptions (for which cf. §4.3.2.1, §4.3.3.1).[58]

**BACKPACK**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *bretella* | 13 | 42 | *laccio* | 6 | 24 |
| *tasca* | 9 | 29 | *manico* | 5 | 20 |
| *cerniera* | 2 | 6.6 | *bretella* | 2 | 8 |
| *carrellino* | 1 | 3.2 | *cinghia* | 2 | 8 |
| *chiusura* | 1 | 3.2 | *maniglia* | 2 | 8 |
| *elastico* | 1 | 3.2 | *attacco* | 1 | 4 |
| *fibbia* | 1 | 3.2 | *attaccatura* | 1 | 4 |
| *gancio* | 1 | 3.2 | *bracciolo* | 1 | 4 |
| *sacca* | 1 | 3.2 | *aggancio* | 1 | 4 |
| *scompartimento* | 1 | 3.2 | *fascetta* | 1 | 4 |
| | | | *fibbia* | 1 | 4 |
| | | | *gancetto* | 1 | 4 |
| | | | *tracolla* | 1 | 4 |
| tot. meronyms | 31 | | tot. meronyms | 25 | |

The most frequent meronym produced in the feature listing task is *bretella* ("strap", 13 occurrences), which is a property related to the practical use of a backpack: the straps rest on the shoulders and allow a person to carry the bag. Since they were visible in the picture of the backpack adopted in the action description task, they were sometimes chosen as target of the grasp, although referred to with different words (*bretella*, *tracolla*, *bracciolo*, *cinghia*). In the semantic feature list, besides the sack (*sacca*), which constitutes the most important part of the bag related to its function, in which people can put things, other parts that are typically found in backpacks occur frequently, such as various kinds of fastenings for closing the bag (*cerniera*, *fibbia*, *elastico*, *chiusura*, *gancio*), pockets and compartments (*tasca*, *scompartimento*), and,

[58] To allow a more easy comparison with the normalised feature norms, all target-related words have been reduced to their singular form. For the stimuli that in the action description task were presented in different orientation condition, only the rightward condition is considered (in order not to have duplicated data for some stimuli).

finally, also the trolley with which backpacks, primarily school bags, are sometimes provided (*carrellino*).

In the action description task, such meronyms are almost never found, because most target-related words variously refer to the top handle of the object (*laccio*, *cinghia*, *attacco*, *attaccatura*, *aggancio*, *fascetta*, *fibbia*, *gancetto*), which is explicitly called "handle" in only seven cases (*manico*, *maniglia*). This part is not directly involved in the main function of the backpack, i.e. that of serving as a container (cf. *sacca*, *tasca*, *scompartimento*), nor to the way in which such bags are usually worn (i.e. on the shoulders). The top handle is only used for short transfers and, moreover, this action could also be performed by using the shoulder straps.

**CHAIR**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *schienale* | 18 | 30.5 | *schienale* | 12 | 48 |
| *gamba* | 17 | 28.7 | *spalliera* | 3 | 12 |
| *seduta* | 8 | 13.6 | *cuscino* | 2 | 8 |
| *bracciolo* | 7 | 11.9 | *asta* | 1 | 4 |
| *ruota* | 5 | 8.5 | *buco* | 1 | 4 |
| *cuscino* | 2 | 3.4 | *sedile* | 1 | 4 |
| *sedile* | 2 | 3.4 | *gamba* | 1 | 4 |
| | | | *seduta* | 1 | 4 |
| | | | *spallina* | 1 | 4 |
| | | | *spazio* | 1 | 4 |
| | | | *stecca* | 1 | 4 |
| tot. meronyms | 59 | | tot. meronyms | 25 | |

For the chair, in both tasks, the most frequently evoked part is *schienale*, "backrest", especially in the action description experiment (where also *spalliera*, with a very similar meaning, also occurs). This part is salient both for the function of the chair (together with the seat and the legs, it sustains a person) and for the action of grasping, because it is the most ready-to-hand part. Also *cuscino*, *seduta*, and *sedile*, which refer to the seat of the chair or to the seat cushion, are shared meronyms between the two groups of features, even if they are more frequent in the property generation task.

However, despite these similarities, the frequency values reported for other meronyms differ greatly. In particular, the legs of the chair (*gambe*) are mentioned 17 times as a relevant property of the object (e.g. "has four legs"), but obviously they do not afford an easy grasp, because they are farther from the hands. Moreover, *ruota* ("wheel") and *bracciolo* ("armrest") occur respectively five and seven times as conceptual semantic features related to the chair

(but the particular chair presented as stimulus in the task did not have either wheels nor armrests). On the contrary, when referring to the target of the grasp, participants often named meronyms of the backrest, such as its vertical slats (*asta*, *spallina*, *stecca*), or the holes between them (*buco*, *spazio*). These parts are never mentioned in the feature listing task, in which only a reference to the backrest as a whole is found.

**COFFEE CUP**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 15 | 100.0 | *manico* | 2 | 100.0 |
| tot. meronyms | 15 | | tot. meronyms | 2 | |

The only meronym named for the coffee cup is the handle, both in the action description and in the property generation task. However, there is an evident difference in the frequency values scored in the two experiments, which is due to the fact that, in the first case, the object-stimulus was overturned, therefore the handle was rarely chosen as the target of the grasp (as already discussed in the previous chapters).

**CUP**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 17 | 100 | *manico* | 20 | 95.3 |
| | | | *occhiello* | 1 | 4.7 |
| tot. meronyms | 17 | | tot. meronyms | 21 | |

Interestingly, the data from the two tasks are quite similar for the cup. More than half of the participants named the handle in the property generation task, as well as for the tea-cup presented in the action description experiment (in which also *occhiello* appears, always referring to the handle).

**GIRL BAG**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *tracolla* | 12 | 36.5 | *manico* | 30 | 91 |
| *manico* | 11 | 33.3 | *bretella* | 1 | 3 |
| *cerniera* | 2 | 6.1 | *maniglia* | 1 | 3 |

| | | | | | |
|---|---|---|---|---|---|
| *tasca* | 2 | 6.1 | *tracolla* | 1 | 3 |
| *borchia* | 1 | 3 | | | |
| *bottone* | 1 | 3 | | | |
| *fibbia* | 1 | 3 | | | |
| *laccio* | 1 | 3 | | | |
| *maniglia* | 1 | 3 | | | |
| *strass* | 1 | 3 | | | |
| tot. meronyms | 33 | | tot. meronyms | 33 | |

For the girl bag, we have already observed in the last chapter that the handles are the only parts indicated as possible target of the grasp (*manico*, *maniglia*, "handle", *bretella*, *tracolla*, "strap"). *Manico* is also the most frequently named meronym for the bag in the property generation task, together with *tracolla*, not only because they constitute the parts with which agents usually interact when using a bag (carrying it by its handles or its strap), but also because they are the parts most commonly shared among different types of bags. Indeed, other characteristics, such as the presence of either a zip, a clasp, or a button fastener (*cerniera*, *fibbia*, *bottone*), of one or more pockets (*tasca*), or of various kinds of ornaments on the bag, such as rhinestones and studs (*strass*, *borchia*), are rather variable among different exemplars of girl bags and are rarely mentioned.

**HAIRDRYER**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *ventola* | 6 | 31.5 | *manico* | 19 | 76 |
| *manico* | 4 | 21 | *impugnatura* | 6 | 24 |
| *pulsante* | 2 | 10.4 | | | |
| *beccuccio* | 1 | 5.3 | | | |
| *comando* | 1 | 5.3 | | | |
| *resistore* | 1 | 5.3 | | | |
| *sifone* | 1 | 5.3 | | | |
| *spina* | 1 | 5.3 | | | |
| *tasto* | 1 | 5.3 | | | |
| *testa* | 1 | 5.3 | | | |
| tot. meronyms | 19 | | tot. meronyms | 25 | |

The object's parts found in the list of feature norms produced for the hairdryer are very different from those chosen as possible targets of the grasp. Only the handle is considered to be a graspable part (*manico*, *impugnatura*, tot. 25 occurrences), whereas *manico* is attested only four times in the norms. On the contrary, in the property generation task, other parts of the object are named, such as *ventola*, "fan", which is even more frequent than *manico*. The

features *testa*, *sifone*, *beccuccio* probably refer to the part of the object from which it blows the air, which, interestingly could be grasped, but is actually never mentioned in the action description task, because it does not correspond to the usual way in which a hairdryer is held. The remaining properties are *pulsante*, *comando*, *tasto* (indicating the buttons), and *spina* ("plug"), all referring to parts more related to the use-plan of a hairdryer, whereas *resistore* ("head") is an internal, non-visible, and obviously non-directly graspable part.

**JUG**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 11 | 55 | *manico* | 23 | 92 |
| *beccuccio* | 5 | 25 | *impugnatura* | 1 | 4 |
| *collo* | 1 | 5 | *occhiello* | 1 | 4 |
| *imboccatura* | 1 | 5 | | | |
| *pancia* | 1 | 5 | | | |
| *tappo* | 1 | 5 | | | |
| tot. meronyms | 20 | | tot. meronyms | 25 | |

In both tasks (but especially in the grasp description one), the part most frequently named is the handle of the jug (*manico*). Among the other meronyms listed as semantic features, informants indicated the neck, the mouth, the spout and the belly (*collo*, *imboccatura*, *beccuccio*, *pancia*), which are constituent parts of most jugs, as well as the stopper that is present in some types of jugs (*tappo*). None of these parts is explicitly mentioned as a possible target of a grasp: the protruding handle is strongly preferred. The total frequency values are similar, although meronyms are slightly more frequent in the action description task.

**LADLE**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 6 | 100 | *manico* | 24 | 88.9 |
| | | | *impugnatura* | 2 | 7.4 |
| | | | *cucchiaio* | 1 | 3.7 |
| tot. meronyms | 6 | | tot. meronyms | 27 | |

In both tasks, the handle is the part of the object most frequently named for the ladle. However, only six informants in the property generation task named it, whereas it occurs 26 times in the action description task. In the latter case, also the concave part of the ladle is mentioned (*cucchiaio*, "spoon"), involved in a two-handed grasp directed to the two

extremities of the object. Interestingly, this part is the one most related to object's function (holding, transporting, stirring, and serving liquids), but it is never mentioned in the feature listing test.

**MICROPHONE**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *filo* | 4 | 57.1 | *manico* | 8 | 53.3 |
| *antenna* | 1 | 14.3 | *impugnatura* | 5 | 33.3 |
| *magnete* | 1 | 14.3 | *gambo* | 1 | 6.7 |
| *tasto* | 1 | 14.3 | *tasto* | 1 | 6.7 |
| tot. meronyms | 7 | | tot. meronyms | 15 | |

The meronyms produced in the two experiments for the microphone are very different. There is only one shared lemma (*tasto*, "button"). For what regards those produced in the feature listing task, *filo* ("cord") and *antenna* (with the same meaning as in English) are optional parts of a microphone, and were not present in the picture representing a microphone in the second task; *magnete* ("magnet") is an internal, non-visible part that obviously could not be directly grasped. In the action description task, the most frequently named meronyms are *manico* and *impugnatura* (*gambo* denotes the same part), and *tasto* occurs only in a phrase in which the target of the grasp was the handle (the part where there is the button).

**PACIFIER**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 3 | 60 | *manico* | 10 | 58.8 |
| *anello* | 1 | 20 | *impugnatura* | 2 | 11.7 |
| *tettarella* | 1 | 20 | *anello* | 1 | 5.9 |
| | | | *gancetto* | 1 | 5.9 |
| | | | *laccio* | 1 | 5.9 |
| | | | *occhiellino* | 1 | 5.9 |
| | | | *occhiello* | 1 | 5.9 |
| | | | | | |
| tot. meronyms | 5 | | tot. meronyms | 17 | |

The meronyms produced for the pacifier in the property generation task are very few. Interestingly, the handle is mentioned more than the teat, which is the part more related to the object's function. The handle is also the only part considered a possible target of the grasp (the teat is excluded for hygienic reasons), but it is mentioned in very different way,

although the most frequent lemma is *manico*, together with its near-synonym *impugnatura*. In particular, we notice the presence of the diminutives *occhiellino*, *gancetto*.

**RUBBER BOAT**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *motore* | 2 | 33.2 | *corda* | 5 | 26.3 |
| *appoggio* | 1 | 16.7 | *manico* | 5 | 26.3 |
| *camera_d_aria* | 1 | 16.7 | *maniglia* | 3 | 15.7 |
| *tubolare* | 1 | 16.7 | *filo* | 2 | 10.5 |
| *valvola* | 1 | 16.7 | *cordino* | 1 | 5.3 |
| | | | *cordoncino* | 1 | 5.3 |
| | | | *elastico* | 1 | 5.3 |
| | | | *passantina* | 1 | 5.3 |
| tot. meronyms | 6 | | tot. meronyms | 19 | |

The meronyms produced for the rubber boat in the property generation task are very few. Most of them refer to features that are proper only of some specific boats: *motore*, and *valvola*, indicate the presence of an engine together with its accessories; *appoggio* denotes the seat. Only two meronyms are related to the flexible inflated tubes, which are typical of all inflatable boats (*camera d'aria*, *tubolare*); however, these words do not occur in the grasp descriptions: we observed that, when the grasp is directed to such parts, informants avoid technical terms and use words related to the spatial domain (e.g. *lati*, "sides"; cf. §4.3.2.1).

None of the properties produced during the feature listing is mentioned as a possible target of the grasp. In the action description task, the handles are indicated eight times (*manico*, *maniglia*), whereas 10 informants variously referred to the rope surrounding the boat (*corda*/*cordino*/*cordoncino*, *filo*, *elastico*), or to the rings through which the rope is inserted (*passantina*). Not all inflatables are provided with these object's parts, but they are quite common, and they were present in the picture constituting the visual stimulus.

**SWORD**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *lama* | 13 | 46.4 | *manico* | 15 | 50 |
| *manico* | 7 | 25 | *impugnatura* | 11 | 36.7 |
| *impugnatura* | 3 | 10.7 | *elsa* | 3 | 10 |
| *punta* | 3 | 10.7 | *punta* | 1 | 3.3 |
| *elsa* | 2 | 7.2 | | | |
| tot. meronyms | 28 | | tot. meronyms | 30 | |

For the sword, there is an evident difference among the meronyms produced both in the property generation and in the grasp description tasks. In the first case, 16 meronyms refer to the blade (*lama*, or, specifically, its extremity, *punta*), whereas 12 features regard the handle (*manico*, *impugnatura*, *elsa*). The high frequency of words related to the blade can be explained considering that it is the part of the object more related to the object's function, whereas the handle is strictly connected to the object's use (the hilt is specifically designed to control the blade, i.e. to use the sword). Obviously, between these two main parts constituting the object, the one that affords the safest grasp is the hilt; therefore, *manico*, *elsa*, and *impugnatura* are much more frequent in the action description task, with respect to *punta*, which occurs only once.

**TROLLEY**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *ruota* | 22 | 46.9 | *manico* | 20 | 80 |
| *manico* | 15 | 31.9 | *maniglia* | 2 | 8 |
| *cerniera* | 5 | 10.6 | *cinghietta* | 1 | 4 |
| *rotella* | 3 | 6.4 | *fascetta* | 1 | 4 |
| *fodera* | 1 | 2.1 | *pezzettino* | 1 | 4 |
| *lucchetto* | 1 | 2.1 | | | |
| tot. | 47 | | | 25 | |

Considering semantic norms, we observe that the most frequent feature produced by the informants is *ruota/rotella* ("wheel/small wheel"), which represents the property that distinguishes a trolley from other suitcases and is strictly related to its use. However, the feature "has a handle" is also produced by a half of the informants, and *manico*, *maniglia* are the object's parts most frequently referred to as the target of the grasp.

Other meronyms named in the action description task all refer to the top handle as well (*cinghietta*, *fascetta*, *pezzettino*). Object's parts produced in the feature listing are not the same as those named in the grasp descriptions: this is true not only for the wheels, which are too low to afford an easy grasp, but also for the zip, the lining, or the lock (*cerniera*, *fodera*, *lucchetto*). We should notice that such parts do not afford the grasp; in particular, *fodera* ("lining") is an internal and, therefore, non-visible part.

**UMBRELLA**

| Property generation task | | | Action description task | | |
|---|---|---|---|---|---|
| meronym | freq. | % | meronym | freq. | % |
| *manico* | 14 | 41.2 | *manico* | 18 | 69.3 |
| *stecca* | 8 | 23.5 | *impugnatura* | 4 | 15.5 |
| *asta* | 4 | 11.9 | *tela* | 1 | 3.8 |
| *punta* | 2 | 5.9 | *collino* | 1 | 3.8 |
| *telo* | 2 | 5.9 | *fusto* | 1 | 3.8 |
| *bottone* | 1 | 2.9 | *corpo* | 1 | 3.8 |
| *levetta* | 1 | 2.9 | | | |
| *pulsante* | 1 | 2.9 | | | |
| *raggiera* | 1 | 2.9 | | | |
| tot. meronyms | 34 | | tot. meronyms | 26 | |

For the umbrella, the most frequently named part in both tasks is the handle, but especially in the action description one (22 occurrences of *manico* or *impugnatura*). Other possible targets of the grasp only regard the body of the object (*corpo*, *fusto*); in particular, its clutched part (*collino*), or the cloth that covers it (*tela*). In the property generation task, we observe that, after the handle, the most frequent meronyms denote structural parts, such as the ribs, the shaft, the spoke, and the cover (*stecca*, *asta*, *raggiera*, *telo*); otherwise, subjects named parts related to the use-plan of an umbrella, such as the button used to open it (*bottone*, *levetta*, *pulsante*). Such parts were not visible in the object presented during the action description task, and were never mentioned.

## 5.5 Comparison between the two tasks

We can now turn to a more general comparison of the meronyms produced in the property generation task and in the grasp description task.

### 5.5.1 Intra-category diversity in meronyms' frequencies

First of all, we should consider the frequency of meronyms produced in the two tasks.

The total number of meronyms provided by 30 informants is 315 in the grasp description task (14 stimuli; mean=22.5 meronyms per stimulus; SD=7.33) and 327 in the property generation task (14 stimuli; mean=23.3 meronyms per stimulus; SD=15.73).

However, there are some evident and significant differences among the single stimuli, related to frequency data (cf. Table 5.3):

| Stimulus | Property generation task | Grasp description task |
|---|---|---|
| Backpack | **31** | 25 |
| Chair | **59** | 25 |
| Coffee cup | **15** | 2 |
| Cup | 17 | **21** |
| Girl bag | **33** | **33** |
| Hairdryer | 19 | **25** |
| Jug | 20 | **25** |
| Ladle | 6 | **27** |
| Microphone | 7 | **15** |
| Pacifier | 5 | **17** |
| Rubber boat | 6 | **19** |
| Sword | 28 | **30** |
| Trolley | **47** | 25 |
| Umbrella | **34** | 26 |
| Tot. | **327** | 315 |

Table 5.3. Total number of meronyms (tokens) produced in the property generation task and in the grasp description task for the artefacts with affording parts. The highest value in each row is in bold type.

There are evident differences among the number of meronyms produced for each stimulus in the two different tasks. A chi-square test conducted on this data yields a highly significant result ($\chi^2(13, N=642)=63.375$, $p<0.0001$), and the analysis of standardised residuals reveal that the stimuli for which the observed values differ more from the expected ones (i.e. std. residuals are smaller/larger than $\pm 1.96$; cf. footnote 27, p. 80) are the chair, the coffee cup, and the ladle (with relation to the frequencies observed for both tasks). For the chair, we have already noticed that the main difference between the two tasks emerged for the meronym *gamba*, "leg", produced 17 times in the property generation task, but only once in the grasp descriptions, and for the features *seduta*/*sedile* (both referred to the seat), which occur 10 times in the feature list, but only two times in the action description task. The coffee cup is characterised by a very low number of meronyms produced with reference to the target of the grasp, with comparison to those listed in the second experiment, and this surely depends on the fact that it was presented overturned, i.e. with a non-canonical orientation (cf. the discussion in §4.3.3.1). For the ladle, instead, informants produced a very low number of meronymic features in the property generation task (there are only six references to its

handle), whereas almost all informants (27 out of 30) named the handle as the target of the grasp.

### 5.5.2 Differences between the meronyms produced in the two tasks

In total, 83 distinct meronyms are produced in the two tasks: 45.8% of them (38) are named only in the property generation task; 28.9% (24) occur only in the grasp description task; and the remaining 25.3% (21) are produced in both tasks (Fig. 5.6).

1. Non-shared meronyms - produced only in the property generation task:
   *antenna*; *appoggio*; *beccuccio*; *borchia*; *bottone*; *camera_d_aria*; *carrellino*; *cerniera*; *chiusura*; *collo*; *comando*; *fodera*; *gancio*; *imboccatura*; *lama*; *levetta*; *lucchetto*; *magnete*; *motore*; *pancia*; *pulsante*; *raggiera*; *resistore*; *rotella*; *ruota*; *sacca*; *scompartimento*; *sifone*; *spina*; *strass*; *tappo*; *tasca*; *telo*; *testa*; *tettarella*; *tubolare*; *valvola*; *ventola*.

2. Non-shared meronyms - produced only in the grasp description task:
   *aggancio*; *attaccatura*; *attacco*; *buco*; *cinghia*; *cinghietta*; *collino*; *corda*; *cordino*; *cordoncino*; *corpo*; *cucchiaio*; *fascetta*; *fusto*; *gambo*; *gancetto*; *occhiellino*; *occhiello*; *passantina*; *pezzettino*; *spalliera*; *spallina*; *spazio*; *tela*.

3. Shared meronyms - produced in both tasks:
   *anello*; *asta*; *bracciolo*; *bretella*; *cuscino*; *elastico*; *elsa*; *fibbia*; *filo*; *gamba*; *impugnatura*; *laccio*; *manico*; *maniglia*; *punta*; *schienale*; *sedile*; *seduta*; *stecca*; *tasto*; *tracolla*.

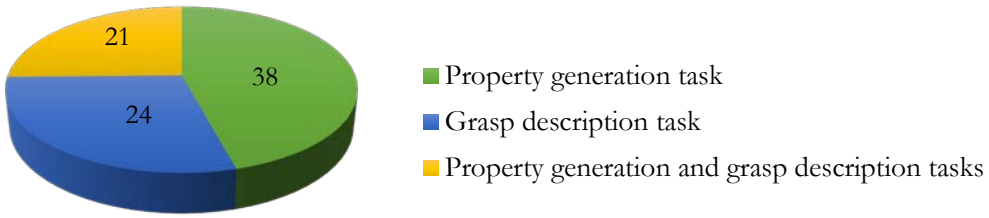



Figure 5.6. Number of distinct meronyms (type frequencies) produced either only in the property generation task, only in the grasp description task, or in both.

Therefore, for what regards the type frequency of meronyms, there is an evident difference in the production of distinct meronyms in the two experiments: a large part of them (74.7%) are produced only in one or the other task (Tab.5.4).

Although most non-shared features are produced in the property generation task (38 vs. 24 types), there is no significant difference between the two groups of meronyms derived from either the property generation or the grasp description tasks, for what regards the number of shared or non-shared types ($\chi^2(1, N=104)=1.300$; $p>0.05$).

| | Non-shared | Shared | Tot. |
|---|---|---|---|
| Property generation task | 38 | 21 | 59 |
| Grasp description task | 24 | 21 | 45 |

Table 5.4. Number of distinct meronyms (types frequencies) either shared or non-shared between the two tasks.

#### 5.5.2.1 Meronyms produced in the property generation task

Meronyms produced in the property generation task are 327, corresponding to a total number of 59 distinct types.

Of these 59 types, as already stated, 38 distinct features (64.4% of meronymic feature types) occur *only* in the property generation task. However, these 38 types correspond to only 111 tokens, i.e. 33.9% over the total number of meronymic features produced (TTR[59]=0.34). As Table 5.5 shows, the majority of the meronyms occurring only in the property generation task have less than five occurrences in the feature list (32 distinct features over 38), and, in particular, 27 features have just one occurrence. Only six features are produced more than five times: among these, we find *cerniera*, "zip" (nine occurrences: two for the girl bag, two for the backpack, and five for the trolley), *tasca*, "pocket" (11 occurrences: two for the girl bag, nine for the backpack), *lama*, "blade" (13 occurrences for the sword), and *ruota*, "wheel" (27 occurrences: five for the chair, 22 for the trolley). These are all objects' parts that do not afford grasping (either because they are not ready-to-hand, such as the wheels, or because

[59] Type-token ratio (henceforth, TTR) is the ratio obtained by dividing the types (i.e. the total number of different meronyms produced in each group of meronyms considered) by its tokens (the total number of meronyms produced), and may range between a theoretical 0 and 1. A high TTR value indicates a high degree of lexical variation among meronyms, whereas a low TTR value means the opposite.

they do not afford a safe grasp, such as the blade), but are relevant for the function of the object and for the events they are usually involved in.

| Meronym | Freq. | % | Meronym | Freq. | % | Meronym | Freq. | % |
|---|---|---|---|---|---|---|---|---|
| *ruota* | 27 | 8.3 | *borchia* | 1 | 0.3 | *raggiera* | 1 | 0.3 |
| *lama* | 13 | 4 | *camera_d_aria* | 1 | 0.3 | *resistore* | 1 | 0.3 |
| *tasca* | 11 | 3.5 | *carrellino* | 1 | 0.3 | *sacca* | 1 | 0.3 |
| *cerniera* | 9 | 2.8 | *chiusura* | 1 | 0.3 | *scompartimento* | 1 | 0.3 |
| *beccuccio* | 6 | 1.8 | *collo* | 1 | 0.3 | *sifone* | 1 | 0.3 |
| *ventola* | 6 | 1.8 | *comando* | 1 | 0.3 | *spina* | 1 | 0.3 |
| *pulsante* | 3 | 0.9 | *fodera* | 1 | 0.3 | *strass* | 1 | 0.3 |
| *rotella* | 3 | 0.9 | *gancio* | 1 | 0.3 | *tappo* | 1 | 0.3 |
| *bottone* | 2 | 0.6 | *imboccatura* | 1 | 0.3 | *testa* | 1 | 0.3 |
| *motore* | 2 | 0.6 | *levetta* | 1 | 0.3 | *tettarella* | 1 | 0.3 |
| *telo* | 2 | 0.6 | *lucchetto* | 1 | 0.3 | *tubolare* | 1 | 0.3 |
| *antenna* | 1 | 0.3 | *magnete* | 1 | 0.3 | *valvola* | 1 | 0.3 |
| *appoggio* | 1 | 0.3 | *pancia* | 1 | 0.3 | | | |

Table 5.5. Meronyms produced only in the feature listing task, which do not occur in the grasp description task (tot. 38 types, 111 tokens). Frequency data and percentages over the total number of 327 meronyms produced by 30 informants.

Moreover, it is interesting to note that, even in the list of meronyms produced in the property generation task alone, most features produced regard visible parts of the entities denoted by the word stimuli. Only some meronyms refer to non-visible parts. In particular, *fodera* (produced for the trolley), *tasca*, *scompartimento* (produced for the backpack), and *raggiera* (produced for the umbrella), denote objects' parts that are visible only when the objects are opened (but consider that *tasca* may refer both to internal and external pockets), whereas *magnete*, *resistore*, and maybe also *valvola* (produced for the microphone, the hairdryer, and the rubber boat respectively) refer to objects' non-directly visible parts, which become visible only if the artefacts are disassembled. These data comply with what Wu and Barsalou (2009: 185) notice with relation to property generation:

> When participants receive a noun or noun phrase, they construct a simulation to represent it, scan across the simulation, and describe properties perceived in the simulation. Because simulations are scanned and described in this manner, unoccluded properties are described relatively often, whereas occluded properties are described less.

Considering now the 21 lemmas found also in the grasp descriptions, they correspond to 216 tokens, i.e. to the 66.1% of the total number of meronyms produced in the property generation task (TTR=0.1). But all features are produced less than 20 times (in particular, 12 types are attested less than five times in the feature list), except one. Indeed, *manico*, "handle", is the most frequent meronym: it occurs 103 times and constitutes 31.5% of the total number of meronymic features produced in the property generation task. Other features are produced much more rarely, such as *schienale*, "backrest" and *gamba*, "leg", referring to the chair (18 and 17 occurrences), as well as *bretella* and *tracolla* ("strap", "shoulder strap") produced, respectively, for the backpack (13 occurrences) and the girl bag (12 occurrences).

| Meronym | Freq. | % | Meronym | Freq. | % | Meronym | Freq. | % |
|---|---|---|---|---|---|---|---|---|
| *manico* | 103 | 31.5 | *bracciolo* | 7 | 2.2 | *tasto* | 2 | 0.6 |
| *schienale* | 18 | 5.5 | *punta* | 5 | 1.5 | *cuscino* | 2 | 0.6 |
| *gamba* | 17 | 5.2 | *asta* | 4 | 1.2 | *elsa* | 2 | 0.6 |
| *bretella* | 13 | 4.0 | *filo* | 4 | 1.2 | *anello* | 1 | 0.3 |
| *tracolla* | 12 | 3.7 | *impugnatura* | 3 | 0.9 | *elastico* | 1 | 0.3 |
| *seduta* | 8 | 2.5 | *fibbia* | 2 | 0.6 | *laccio* | 1 | 0.3 |
| *stecca* | 8 | 2.5 | *sedile* | 2 | 0.6 | *maniglia* | 1 | 0.3 |

Table 5.6. Meronyms produced in the feature listing task that occur also in the grasp description task (tot. 21 types, 216 tokens). Frequency data and percentages over the total number of 327 meronyms produced by 30 informants.

Therefore, with regards to the features produced in the property generation task, we can conclude that the types of meronyms found *only* in the feature listing task are more numerous than those produced *also* in the other task (38 vs. 21); however, the majority of the tokens (216 vs. 111) belong to the group of the shared meronyms (cf. Table 5.7). The most frequent meronyms are derived from a few lemmas and are shared between the two tasks. In particular, the feature "handle" (*manico*) stands out for its very high frequency.

| | Non-shared | Shared |
|---|---|---|
| Meronyms (types) | 38 | 21 |
| Meronyms (tokens) | 111 | 216 |
| TTR | 0.34 | 0.1 |

Table 5.7. Types, tokens, and type-token ratios of the meronyms produced in the feature listing task (14 stimuli), either shared or non-shared with the meronyms produced in the grasp description task.

#### 5.5.2.2 Meronyms produced in the grasp description task

Meronyms produced in the grasp description task are 315. They correspond to a total number of 45 distinct types, 24 of which (53.3%) occur *only* in the grasp description task.

However, these 24 meronyms occur only 35 times (TTR=0.7), and therefore regard only 11.1% of the total number of meronyms produced in the grasp description task. As shown in Table 5.8, most of these meronyms (18) are represented only once, and the token frequency never exceeds five occurrences. We notice that there is a relatively high number of diminutives (such as *fascetta*, *cinghietta*, *cordoncino*, *passantina*, *pezzettino*, *collino*) that are never produced in the property generation task. This diversity can be explained considering that diminutives typically occur much more freely in the oral than in the written discourse (e.g. Dressler and Barbaresi 1994): in the grasp description task, meronyms were orally produced in an interview, whereas in the property generation task they were written on a questionnaire in a more formal and controlled situation.

| Meronym | Freq. | % | Meronym | Freq. | % | Meronym | Freq. | % |
|---|---|---|---|---|---|---|---|---|
| *corda* | 5 | 1.6 | *attacco* | 1 | 0.3 | *fusto* | 1 | 0.3 |
| *occhiello* | 3 | 1.0 | *buco* | 1 | 0.3 | *gambo* | 1 | 0.3 |
| *spalliera* | 3 | 1.0 | *cinghietta* | 1 | 0.3 | *occhiellino* | 1 | 0.3 |
| *cinghia* | 2 | 0.7 | *collino* | 1 | 0.3 | *passantina* | 1 | 0.3 |
| *fascetta* | 2 | 0.7 | *cordino* | 1 | 0.3 | *pezzettino* | 1 | 0.3 |
| *gancetto* | 2 | 0.7 | *cordoncino* | 1 | 0.3 | *spallina* | 1 | 0.3 |
| *aggancio* | 1 | 0.3 | *corpo* | 1 | 0.3 | *spazio* | 1 | 0.3 |
| *attaccatura* | 1 | 0.3 | *cucchiaio* | 1 | 0.3 | *tela* | 1 | 0.3 |

Table 5.8. Meronyms produced in the grasp description task that do not occur in the feature listing task (tot. 24 types, 35 tokens). Frequency data and percentages over the total number of 315 meronyms produced by 30 informants.

In contrast with what we observed for the non-shared meronyms, the 21 lemmas shared with the property generation task correspond to 280 tokens, i.e. 88.9% of meronyms produced in this task (TTR=0.08). Also in this case, most types (16), corresponding to 23 tokens, are attested less than five times among the lemmas denoting target-related words (11 occur only once). The vast majority of tokens correspond to only two distinct features, *impugnatura* (31 occurrences, i.e. 9.8% over the total number of meronymic features produced in the grasp description task), and *manico* (199 occurrences, 63.2%).

| Meronym | Freq. | % | Meronym | Freq. | % | Meronym | Freq. | % |
|---|---|---|---|---|---|---|---|---|
| *manico* | 199 | 63.2 | *cuscino* | 2 | 0.7 | *fibbia* | 1 | 0.3 |
| *impugnatura* | 31 | 9.8 | *filo* | 2 | 0.7 | *gamba* | 1 | 0.3 |
| *schienale* | 12 | 3.8 | *tracolla* | 2 | 0.7 | *punta* | 1 | 0.3 |
| *maniglia* | 8 | 2.5 | *anello* | 1 | 0.3 | *sedile* | 1 | 0.3 |
| *laccio* | 7 | 2.2 | *asta* | 1 | 0.3 | *seduta* | 1 | 0.3 |
| *bretella* | 3 | 1.0 | *bracciolo* | 1 | 0.3 | *stecca* | 1 | 0.3 |
| *elsa* | 3 | 1.0 | *elastico* | 1 | 0.3 | *tasto* | 1 | 0.3 |

Table 5.9. Meronyms produced in the grasp description task that also occur in the feature listing task (tot. 21 types, 280 tokens). Frequency data and percentages over the total number of 315 meronyms produced by 30 informants.

For what regards the meronyms produced in the action description task, we can conclude that the meronyms (types) found *only* in the feature listing task are more numerous than those produced also in the other task (24 vs. 21); however, also in this case, the vast majority of tokens (429 vs. 48) are found in the group of the shared meronyms.

The most frequent meronyms are derived from a few lemmas and are shared between the two tasks. In particular, meronyms denoting the "handle" (*manico*, *impugnatura*) are the most frequent ones.

| | Non-shared | Shared |
|---|---|---|
| Meronyms (types) | 24 | 21 |
| Meronyms (tokens) | 35 | 280 |
| TTR | 0.69 | 0.08 |

Table 5.10. Types, tokens, and type-token ratios of the meronyms produced in the grasp description task (14 stimuli), either shared or non-shared with the meronyms produced in the property generation task.

#### 5.5.2.3 Results of the comparison

We observed that the meronyms (tokens) produced in the property generation task were, in mean, 23.3 per stimulus (327 meronyms per 14 stimuli). Now we are able to specify that, in mean, 7.9 meronyms per stimulus emerged in this task alone, whereas 15.4 meronyms per stimulus are found in the other task as well. In the grasp descriptions, informants produced 315 meronyms for 14 stimuli, with a mean of 22.5 meronyms per stimulus; but a mean of

only 2.5 meronyms per stimulus is found only in this task, whereas a mean of 20 meronyms per stimulus are shared between the two tasks.

Therefore, the two groups of meronyms respectively produced in the property generation task and in the grasp description task are different from each other, not in the overall number of meronyms produced, nor in the number of distinct meronyms (types) either shared or non-shared between them (cf. §5.5.2), but with respect to the number of occurrences (tokens) of shared and non-shared meronyms ($\chi^2(1, N=632)=64.621$; $p<0.0001$), as Figure 5.7 shows.

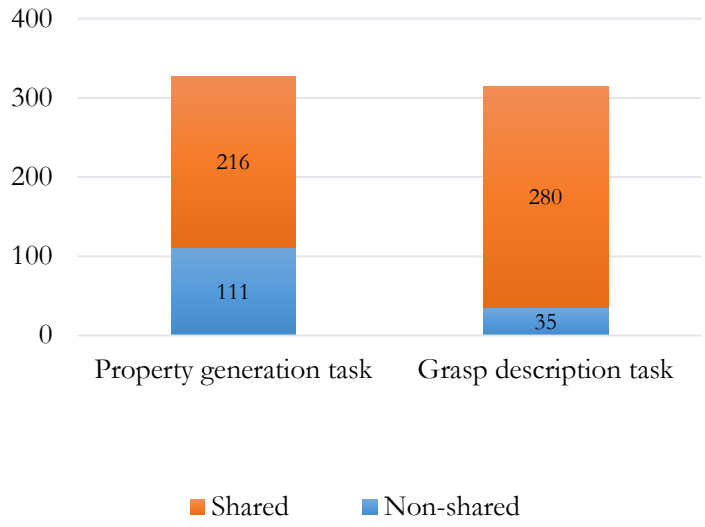



Figure 5.7. Shared vs. non-shared meronyms (tokens) produced in the property generation task (14 stimuli, 327 meronyms) and in the grasp description task (14 stimuli, 315 meronyms).

We observed in the previous paragraphs that in the property generation task there are some features that are often produced for a stimulus, but that are never judged as a possible target of the grasp, such as *lama*, "blade" (produced 13 times for the sword). Instead, in the grasp description task, non-shared meronyms are generally produced less frequently (e.g. *corda*, "rope", produced five times for the rubber boat).

On the other hand, the shared meronyms are much more frequently produced in the grasp description task (e.g. *manico*, "handle", has 199 occurrences) than in the property generation task (in which *manico* has 103 occurrences).

## 5.6 Discussion

The analysis conducted has revealed some differences, as well as some similarities among the meronyms produced with reference to 14 artefacts either in the grasp description task or in the property generation task.

First of all, we have noticed (cf. Fig. 5.6) that there is a large number of distinct meronyms that emerges only in one or in the other task. In part, this is due to an intrinsic limit of this comparison, i.e. to the great difference between the two tasks, not only in the methodology and procedure adopted, but also in the purposes of the tasks. Although the two experiments were conducted on the same number of subjects (with similar age) and were focussed on the same set of items, in the first task participants were visually presented with an image alone (without any written word), representing a particular instance of an object, whereas in the second task they were presented with only a written word (without any image). Moreover, in the action description task, informants were requested to describe (orally) how they would have grasped the objects represented, whereas in the property generation task they were asked to list (in a written form) the features that they judged relevant to describe the meaning of the words (cf. also *infra*). Therefore, it is obvious that in the grasp description task participants could only name those object's parts that were actually perceivable in a given picture. For instance, the particular inflatable presented as an example of a rubber boat had no motor; therefore, the motor could not be named in the descriptions of the grasp. Conversely, the same rubber boat was provided with a rope inserted in small rings, but not all rubber boats have ropes and rings, and such meronyms are not listed among the properties generated in the feature listing task. Nevertheless, as previously discussed, there are some objects' parts that could have been named as the target of the grasp, but in fact they were not: for instance, the neck and the mouth of the vase (*collo*, *imboccatura*).

However, we have also noticed that the meronyms that occur in only one of the two tasks are usually produced very rarely, as the type-token ratios reported in Tables 5.7 and 5.10 show. On the contrary, the largest number of meronyms (in terms of token frequency) produced in the two tasks widely overlaps (tot. 496 tokens for 21 distinct meronyms): most meronyms produced in the property generation task are also those that afford manipulation and can be indicated as the target of the grasp. In particular, the detailed analysis conducted in the last paragraph shows that *manico* is the most frequent meronym occurring in the grasp

description task (199 occurrences), as well as in the property generation task (103 occurrences).

Handles are affording parts, the parts of an object most typically involved in the use-plan of an artefact. We usually (and sometimes necessarily) grasp an artefact's handle before using the object (e.g. a ladle, a cup, a sword, etc.). In this regard, it is worth recalling that a growing body of research demonstrates that words denoting manipulable objects, as well as their visual presentation, automatically activate affordances, intended as the motor information regarding habitual micro-interactions with their referents (cf. Borghi 2005; 2007; cf. §1.4.2). Many artefacts' parts frequently named in the property generation task here described (in particular *manico*, *schienale*) are also frequently named as target of the grasp. Such parts (in particular handles) are usually involved in the habitual micro-interactions that agents carry out with objects, since object grasping is preliminary and necessary to perform the typical actions for which the artefacts have been designed. This is particularly evident for *manico*, but it is true also for the chair, the only object that does not have a handle, for which *schienale* is the meronym more frequently produced in both tasks.

However, if we consider frequency data reported in Tab. 5.6 and Tab. 5.9, it seems evident that the meronym *manico* is much more frequently produced in the grasp description task (63.2% of 315 meronyms) than in the property generation experiment (31.5% over 327 meronyms). Moreover, *manico*, *maniglia*, and *impugnatura* (which occur in both tasks) can be considered as near-synonyms in this specific case, since for the stimuli adopted in the two experiments they all denote the object's handle; therefore, we can even merge their occurrences and compare the number of meronyms denoting an object's handle (restricted to only the three lemmas *manico*, *maniglia*, and *impugnatura*), with any other different meronym:

| | handle | other meronyms | tot. meronyms |
|---|---|---|---|
| Property generation task | 107 | 220 | 327 |
| Grasp description task | 238 | 77 | 315 |

Table 5.11. Frequency of the meronyms *manico*, *maniglia*, and *impugnatura* compared to the frequency of all other meronyms produced either in the property generation or the grasp description task.

Table 5.11 shows an evident and significant difference between the two groups of meronyms, produced either in the property generation task or in the grasp description task,

for what regards the number of occurrences of *manico*, *maniglia* and *impugnatura* compared to that of all other meronyms ($\chi^2(1, N=642)=118.4; p<0.0001$).

This result indicates that the difference in the two tasks has a strong impact on token frequency data. Since in the first experiment the informants were explicitly requested to imagine a particular action performed on artefacts, i.e. a grasp, meronyms production is evidently biased towards affording parts *specifically designed* for the grasping and strongly related to this specific function. The parts with which agents usually interact are, instead, less consistently listed in the property generation task: their important presence, however, confirms that at least a large part of the meronyms produced is strictly linked to the concrete and repeated interactions with objects.

# Chapter VI

# Possible Applications and Future Directions

This last chapter reports two case studies that, although being very different from one another, are strictly related to the study presented in this thesis and are part of the research activity conducted within the ImagAct and ModelAct[60] projects (cf. Introduction).

The first one introduces a possible application of the notion of affordances in computational linguistics; in particular, it explores the possibility of annotating the information related to the affordances for grasping in a corpus in order to help with word sense disambiguation (§6.1). The second study is more relevant to cognitive science and offers a more in-depth investigation of a topic introduced in Chapter III, i.e. the effect of the object's orientation on the grasp descriptions. It also describes a possible exploitation of the data collected during the interviews also for purposes other than those for which they were originally collected: in particular, gesture analysis (§6.2). Finally, some possible future directions will be proposed (§6.3).

[60] For all information about the project, visit the website http://modelact.lablita.it/.

## 6.1 Affordances and word sense disambiguation: a case study

The first chapter of this thesis reviewed many behavioural and brain imaging studies that demonstrate that, at least for the specific field of grasping and manipulation, the simple perception of an object is able to automatically activate affordances, intended as sensory-motor simulations of the typical micro-interactions between an agent/perceiver and an object (e.g. Ellis and Tucker 2000; Grèzes, Tucker et al. 2003). The activation of such sensory-motor simulations is modulated by a number of factors: in particular, by objects' dimension and shape, and by the presence of an affording part. We should recall that Borghi and Riggio (2009) call *stable affordances* the affordances related to invariant features or properties of objects (such as dimension and shape), incorporated into an object's representation, on which agents build a motor prototype. From a similar point of view, it is worth noting that Gibson defines an affordance also as «an invariant combination of variables» (Gibson 1979: 134). As stated by Michaels (2003: 146), actions are afforded by a «multidimensional compound» of objects' properties.

By knowing what an agent can do with an object, according to the object's physical, perceivable (stable) properties, it may be possible to distinguish the types of events in which the object can be involved ("doable" actions), as well as those in which it cannot be involved ("non-doable" actions). Therefore, the opportunity to annotate in a corpus the information related to affordances, intended as the possibilities for action offered by objects' visually perceivable properties, could be useful in dealing with problems regarding polysemy and word sense disambiguation.

A first case study (cf. De Felice 2014b)[61] explores this issue by annotating the information related to the affordance of grasping in a corpus specifically focussed on Italian grasp verbs. The corpus was annotated at both a "pragmatic" level (with features related to objects' affordances) and at a semantic level (with semantic features related to objects' superordinate categories), and underwent a classification experiment in order to test which kind of information best distinguishes concrete from metaphorical uses of a verb.

[61] The contribution *"Possibilities of action" in language: affordances and verbal polysemy* was presented at the 7th Conference of the Italian Network of Doctoral Schools for Cognitive Sciences (CODISCO 2013), held in Noto from 28 to 30 November 2013, and has been published in the journal *Reti, Saperi, Linguaggi. Italian Journal of Cognitive Sciences*.

### 6.1.1 Methods

The experiment was conducted on the corpus developed within the ImagAct project. Summarising what has already been explained in Chapter II (in particular, cf. §2.3.1-2), this corpus[62] was specifically created to study the semantic variation of action verbs: all occurrences of more than 1100 Italian and English action verbs were extracted from pre-existing spoken corpora with their linguistic context; then, each context extracted was standardised and reduced to a very simple sentence in order to create, for each lemma, a list of short sentences (e.g. *John takes the book from the shelf*) that show which kind of events people usually refer to when using action verbs in real conversation. Then, expert annotators assigned a primary or marked value to all standardised instances (Gagliardi 2014: 33): in a primary sentence, the action verb refers to a concrete and physical action (e.g. *John gives Mary the umbrella*), whereas, in a marked one, the verb is used in an abstract, metaphorical, or idiomatic expression (e.g. *John gives me a good idea*). Since the purpose of the ImagAct project was to study the semantic variation of action verbs in their primary usages, only primary standardised sentences of Italian action verbs were additionally annotated with syntactic and semantic features (lemmas, thematic roles, aktionsart value).

In order to conduct the experiment on affordances annotation, the verbs denoting an action of grasping (*acchiappare*, *accogliere*, *afferrare*, *cogliere*, *pigliare*, *prendere*, *raccattare*) were selected among the Italian verbs included in the ImagAct database (almost 600 lemmas). All their primary and marked standardised sentences (tot. 2802; primary=1309; marked=1493) were collected. Since only primary sentences had already been annotated with syntactic information, all marked sentences were also parsed in order to extract direct objects together with their lemmas (tot. direct objects' lemmas: 2779 tokens, 714 types).

Whereas the sentences in which the direct object of a verb of grasp is an abstract entity are clearly always marked (e.g. *Marco prende freddo /prende una decisione / coglie l'occasione*), the instances in which there is a concrete entity involved in the event (2188, i.e. 78% of the total number of 2802 sentences) may have either a marked or a primary value. For instance, *Marco prende il treno* ("Mark catches the train") is a marked sentence, whereas *Marco prende il trenino*

[62] We should recall that what we refer to as the "ImagAct corpus" has many peculiarities; it has been developed for a specific purpose and does not fully satisfy the traditional principles of sampling and representativeness (cf. McEnery and Wilson 2001: 32). On this topic, cf. Gagliardi (2014: 58).

("Mark takes the toy train") is a primary one: both *treno* and *trenino* are artefacts, but only in the last sentence does the verb *prendere* properly mean "to grasp".

The annotation of the corpus was conducted in two phases. First of all, the direct object's **semantic type** was annotated adopting the WordNet SuperSenses categories (Ciaramita and Johnson 2003; Ciaramita and Altun 2006) pertaining to concrete objects: *top* (top ontology elements, as for *element*), *animal* (animals, as for *dog*), *artefact* (man-made objects, as for *pencil*), *body* (body and body parts, as for *hand*), *food* (foods and drinks, as for *bread*, *wine*), *object* (natural objects, as for *stone*), *person* (people, as for *Mark*), *plant* (plants and their parts, as for *flower*, *leaf*), *substance* (substances and materials, as for *water*, *sand*). Additionally, the semantic type of subjects was also annotated, distinguishing only among *humans*, who are the most common subjects both in primary (98%) as well in marked (84%) sentences, *animals* (e.g. *la scimmia prende la banana*), and *inanimates* (e.g. *il camion raccoglie la spazzatura*).

In the second phase, the **kind of grasp** afforded by the objects was manually annotated. The four categories relative to the kind of grasp afforded were created adopting a bottom-up approach, by looking at all the possible objects of primary occurrences of verbs, and by identifying a minimum set of common stable and visually perceivable features among them. Therefore, the categories considered are mainly related to objects' size, constituency, and shape. Each one of them will be briefly presented.

**One-handed grasp**: this kind of grasp is afforded by objects that have no handles or protruding parts, and that can be grasped by using only one hand. The size of two of the object's dimensions (length, width or thickness) usually does not exceed the maximum span of a hand with at least two fingers bent in order to grasp and hold something (e.g. *boccia*, "bowl", *bottiglia*, "bottle", *candela*, "candle", *conchiglia*, "shell", *collana*, "necklace", *molletta*, "clothes peg").

**Two-handed grasp**: this second type of grasp is afforded by objects that have no handles or protruding parts that could attract the grasp and that are usually grasped with two hands (e.g. *asse*, "board", *pallone*, "soccer ball", *pianola*, "player piano", *grammofono*, "gramophone", *computer*).

**Grasp by part**: this third type of grasp is afforded by:

(i) entities that, usually, have a part specifically designed for the grasping (e.g. *coltello*, "knife", *brocca*, "jug", *accetta*, "axe"), which is generally not expressed linguistically in the standardised sentences (*Marco prende la brocca*, and not *Marco prende la brocca dal manico*);

(ii) entities that have a part that, even if it is not specifically designed for this specific purpose, is more suited than others for the grasping thanks to its shape and conformation (a grasp by part may be directed either to objects, such as *sedia*, "chair", or to humans). For their relevance for action, as well as because they are not predetermined, such objects' parts are often explicitly mentioned in the standardised sentences, especially if there are many possible graspable parts in the same entity, as in *John takes Mary by her hand/her leg/her arm*.

**Grasp with instrument-container**: this kind of grasp is afforded mainly by substances and aggregates, which humans in everyday activities usually do not control without using some other object (an instrument, generally a container), because of their fluid consistency and because of the absence of a solid, definite shape contour (e.g. *acqua*, "water", *brodo*, "broth", *farina*, "flour", *crusca*, "bran"). Since the most frequent kind of interaction with these entities is with a recipient, the explicit reference to the container is often left understood. Consider the following examples:

1. *Marco piglia l'acqua dal rubinetto* ("Mark takes the water from the faucet")
2. *Marco prende l'acqua per il cane* ("Mark takes the water for the dog")

In these sentences, it is implicitly understood that the agent uses a container in order to perform the action expressed by the verb (for instance, we could imagine a glass or a bottle for the first instance, a bowl for the second one). Since the most common way in which water is taken from a faucet is, by default, by using a container, the information relative to the instrument with which the action is performed is redundant and therefore not expressed linguistically. In such cases, a sort of affordance-based knowledge about the referent of a lexical item (in this case, relative to the way in which liquids are usually taken) seems to constrain and determine inferences (Attardo 2005: 172). [63] In this regard, the notion of

[63] We just touch on this very intriguing topic, strictly related to the semantics/pragmatics interface, which would be well worth exploring in future research. For this purpose, the ImagAct corpus would prove very useful, especially because standardised sentences derived from the occurrences of grasp verbs often contain

*lexicalised pragmatics*, as discussed by McDonald and Pustejovsky (2014), is a particularly relevant one. [64]

The classification adopted, which is based on a distinction of a one-handed grasp, two-handed grasp, grasp by part, and grasp with instrument-container, complies well with the results of the analysis presented in Chapters III and IV. We observed that the grasp of natural kinds, such as the mandarin and the stone, or objects without affording parts, is mostly described as an undifferentiated unimanual or bimanual grasp, depending on the object's size (the references to the hand are the most frequent, whereas the target is rarely mentioned; e.g. for the mandarin: «questo con una mano, semplicemente»; for the soccer ball: «con due mani»). The grasp described for artefacts with affording parts and humans is mostly directed to entities' parts (the effector of the grasp is rarely mentioned, whereas in most cases the target of the grasp is explicitly named; meronyms are frequently produced; e.g. for the man: «per una mano»; for the girl bag: «dal manico»). Finally, for substances and aggregates, we observed a very high number of explicit references to containers and instruments, at least evoked in similes and analogies (cf. §4.7.1.3).

The 2802 annotated instances were imported as a data frame in the R statistical computing environment[65] and the decision tree predictive model (for which cf. *infra*) was used to investigate which features, among the following variables, are more effective in predicting the primary or marked value of the instances: (i) a subject's (*human/animals/inanimate*) or an object's semantic type (*top*; *animal*; *artifact*; *body*; *food*; *object*; *person*; *plant*; *possession*; *substance*); (ii) the type of grasp afforded by the object (*one-handed grasp*; *two-handed grasp*; *grasp by part*; *grasp*

---

prepositional phrases (for instance, we find also *Marco prende l'olio dal rubinetto con la bottiglia*, in which the instrument is explicitated).

[64] «In a lexicalized grammar, the terminals of the rules are specific words instead of lexical categories such as proper noun or transitive verb. We propose to lexicalize meaning and inference – to establish it directly from the incremental composition of the meaning of the words in a text without an intervening logical form. The meaning of words, phrases, and meaning-bearing constructions is defined in terms of the set of entities, predicates, relations, propositions, or potential inferences they convey. Situations are created dynamically by composing these *packets* of content and inference as the words of a text are scanned. Most packets correspond to small individual categories or inferences, such as the affordances of a cup as a container, or the consequences of a process being canceled» (McDonald and Pustejovsky 2014: 146; original emphasis).

[65] For information and references, visit the website of The R Project for Statistical Computing at http://www.r-project.com.

*with instrument-container*). Data were randomly divided into a training set (70%) and a test set (30%); then, three experiments were conducted (using the rpart package).

### 6.1.2 Results and discussion

The results of the classification experiments conducted show that training the algorithm on semantic variables alone (WordNet SuperSenses), 73.48% of test instances are correctly classified as primary or marked. Considering only the type of grasp afforded, 82.88% of test instances are correctly classified. Combining all features, the predictive model[66] (Fig. 6.1) assigns the correct primary/marked value to 84.3% of test instances (with 95 marked sentences incorrectly classified as primary and 37 primary sentences incorrectly classified as marked).

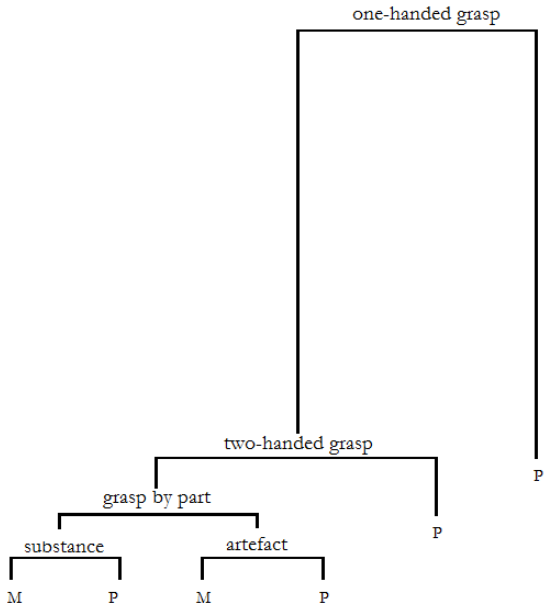



Figure 6.1. The tree predictive model generated in the third classification experiment (M=marked, P=primary).

[66] In this classification model, the nodes of the tree are constituted by the variables used as predictors (one-handed grasp, two-handed grasp, etc.). The possible values of such variables are represented by the branches that connect parent to child nodes. From each node, the left branch corresponds to the 0 value (when the direct object of a sentence has not been annotated with a given feature), whereas the right branch corresponds to the 1 value (when the direct object has been annotated with a given feature). The leaves are labelled with the values predicted for the target variable (the primary or marked value of a sentence).

Most errors are due to ambiguous instances involving concrete entities (affording a unimanual grasp) that allow a primary as well as a marked interpretation. *Marco prende il biglietto del concerto*, for instance, can mean that Mark concretely "takes" a ticket for a concert (primary), but also that Mark "buys" the ticket (marked). Therefore, many instances have been correctly classified by annotators because they could read the original complete sentence extracted from the speech corpora adopted in the project, but their disambiguation is impossible (also for humans) if a wider linguistic context is omitted.

This first case study, although being conducted on a small sample of data annotated by a single person, yields encouraging results and suggests that the pragmatic knowledge related to affordances, i.e. to the habitual kind of interactions that agents have with objects, might be useful in order to decide which type of event is denoted by a given sentence, whether a primary (referred to a concrete action) or a marked one. As a future work, it would be interesting to extend the same annotation scheme to English data (primary and marked standardised sentences containing verbs such as *to take*, *to grasp*, *to pick*), therefore enriching the number of objects that can be involved in the event of grasping and can be annotated with the features related to the kind of grasp afforded.

Since the first objective of the ILC Research Unit in the ModelAct project is to develop efficient strategies for the automatic disambiguation of the ImagAct action types (for which cf. §2.3.1), the role of affordances annotation was further investigated. Therefore, an annotation schema that also includes the type of grasp that objects typically afford was also adopted in Russo, Frontini et al. (2013).[67] This experiment, too, was conducted on the ImagAct corpus: in particular, on all primary instances collected for the Italian verbs roughly corresponding to Levin's HOLD verbs (cf. Levin 1993), i.e. *acchiappare*, *afferrare*, *agguantare*, *pigliare*, *prendere*, *raccattare*, *raccogliere*, *stringere*, *tenere* (for a total of 1419 sentences). The purpose of this experiment was to disambiguate not between marked and primary sentences, but among different action types (therefore this study is conducted only on primary instances), which, for the verbs considered in this experiment, are 29. All direct objects present in the corpus considered were annotated, first of all, with the type of grasp afforded and with their semantic type. Crucially, they were also annotated with information derived from telic relations extracted from SIMPLE, a lexical resource largely based on Pustejovsky's

[67] The contribution has been presented at the 6th International Conference on Generative Approaches to the Lexicon, held in Pisa from 24 to 25 September 2013.

Generative Lexicon theory (cf. Pustejovsky 1991; 1995),[68] by considering their values encoded for the Telic quale (e.g. for *Matteo prende il coltello*, *coltello*, "knife" is used-for *tagliare*, "to cut"), classified according to the verbs' abstract semantic classes considered in SIMPLE (*tagliare* pertains to the cause-constitutive-change verb class). Finally, direct objects' lemmas were also annotated with co-occurrence information extracted from the it-TenTen corpus[69] (by considering the three most salient verbs that have as object the ImagAct nouns, classified, again, according to SIMPLE semantic classes). The classification experiment conducted (using the Support Vector Machine (SMO) classification algorithm implemented in WEKA)[70] revealed that the kind of grasp afforded by objects, considered alone, is slightly less relevant for disambiguating among the 29 different action types (accuracy=76%) with comparison to semantic (accuracy=77%) and qualia information (accuracy=77.4%), whereas co-occurrence information extracted from corpora is the most useful (accuracy=80.5%). Considering all features together, accuracy reaches 81.6% (for details about the experiment and the results, see Russo, Frontini et al. 2013: 74). The affordances annotation proposed is not very helpful in this case, because different types of actions in ImagAct do not always correspond (at least in this class of verbs) to different types of objects, distinguished only by the kind of grasp required by their physical properties (such as dimension and shape).

In conclusion, the two experiments here described draw the attention to the possibility of exploiting affordance-based knowledge in language technologies and lexical resources.[71] This

[68] According to the Generative Lexicon theory, lexical expressions can be represented at different levels, one of which, the qualia structure, consists of four roles (each one encoding separate aspects of the word's meaning): the formal (what an entity is), the constitutive (what an entity is made of), the telic (what the function of an entity is), and the agentive one (how an entity came into being). SIMPLE (Lenci et al. 2000; Ruimy et al. 2003) is based on the notion of Extended Qualia Structure, whereby each of these four roles subsumes a set of semantic relations (for instance, the telic relation, which provides essential knowledge about an object's typical uses, subsumes relations such as *used-for*, *object-of-the-activity*, *used-as*, among others).

[69] It-TenTen is a web corpus of 3.1 billion tokens, accessible at http://www.sketchengine.co.uk/ (cf. Kilgarriff et al. 2014).

[70] http://www.cs.waikato.ac.nz/ml/weka/ (cf. Hall et al. 2009).

[71] For the possibility to acquire information relative to affordances from web corpora, cf. Russo, I., De Felice, I., Frontini, F., Khan, F., and Monachini, M. (2013). (Fore)seeing actions in objects. Acquiring distinctive affordances from language. In *Proceedings of The 10th International Workshop on Natural Language Processing and Cognitive Science - NLPCS 2013* (Marseille, France, 15-17/10/2013), 151-161.

is still an under-investigated issue that may open new and interesting pathways for research and could also have a considerable theoretical impact.

## 6.2 Beyond the linguistic surface of grasp descriptions: language, gestures and objects' orientation

Chapter I presented a body of research demonstrating that the activation of sensory-motor simulations is modulated by a number of factors: in particular, by the possibility for an agent to perceive the affording part of a graspable object, especially when it is spatially aligned with the hand for which brain activity is measured (Grèzes, Tucker et al. 2003; Buccino et al. 2009). Such effect of spatial alignment on affordances activation has also been proven by many behavioural experiments based on compatibility paradigms (Tucker and Ellis 1998; Phillips and Ward 2002; Symes et al. 2007).

In Chapter III (§3.2.3.3), we observed that spatial alignment seems to have an effect on the frequency of linguistic references to the effector or the target of the grasp. Also, the analysis conducted in Chapter IV has highlighted that the presence of spatial terms related to the right or the left hand might be indicative of a spatial compatibility effect of the object's orientation on the type of grasp described, as far as reflected in the linguistic descriptions (§4.7.1.2). However, linguistic data are insufficient to classify the types of grasps actually described by informants. We noticed, for instance, that for many stimuli participants did not make any explicit mention of either the target or the effector of the grasp. Even if it goes beyond the scope of this thesis, there is another important layer that could be added to transcripts annotation and could help in classifying grasps, thus also allowing a comparison with the grasp taxonomies: that of gestural production.

As a growing body of literature suggests, gestures are a mode of expression tightly linked to language and speech (e.g. Holler and Beattie 2003; McNeill 1992).[72] This is also confirmed by neuropsychological evidence for a strong correlation between the severity of aphasia and the severity of impairment in gesture production (Cocks et al. 2013). As stated by Kendon (2004: 2-3, original emphasis):

[72] As happily stated by Guellaï et al. (2014), «human language is a multimodal experience: it is perceived through both ears and eyes».

> As a little reflection shows, the way in which gesture and speech serve as modes of expression is quite different. Speech uses an established vocabulary of lexical forms organized in structures that unfold as a temporal succession, according to rules of syntax. Gesture, on the other hand, especially when used in conjunction with speech, tends not to have these features and is often regarded as expressive because it is depictive or pantomimic. Yet, how can a person, in creating an utterance, *at one and the same time*, use both a language system and depictive and pantomimic actions? As a close examination of the coordination of gesture with speech suggests, these two forms of expression are integrated, produced together under the guidance of a single aim.

Together with the recent growth of interest in gesture, many classifications, transcription conventions, and annotation systems have been proposed (cf. McNeill 1992: 75 ff.; Kendon 2004: 362 ff.; Kipp 2004: 49-72; Bressem et al. 2013).

The brief analysis presented in the following paragraphs explores the effect of object orientation on the grasp types described by the 30 subjects in the action description task. Crucially, the gestural information, retrieved by the video recordings of the interviews, was added to the linguistic information derived from the transcripts (on which the analysis of Chapters III and IV was based). While describing a grasp, the participants usually enacted the action they were describing; therefore, even if they rarely linguistically specified, for instance, whether the effector was the right or the left hand, they often provided this information with gestures.

### 6.2.1 Methods

This case study has been conducted on six visual stimuli representing two artefacts with an affording part, the jug and the tea-cup, in all their orientation conditions (leftward/rightward orientation for both; upright/overturned only for the tea-cup). The 180 grip descriptions provided for these six stimuli were classified according to a series of parameters, considering linguistic as well as gestural production (cf. *infra*).

For what regards the target of the grasp, only grasps that are directed to an object's handle (which is the only part specifically designed for grasping in these objects) and grasps that are directed to another object's part (e.g. grasps from the above or lateral grasps directed to the body of the object) were distinguished. In most cases, this information was provided linguistically (cf. Chapter IV) and informants either referred to the handle explicitly naming

it (1), or by indicating it in terms of spatial relations (2), as the following examples from the leftward-oriented tea-cup show:

(1) *029: dal manico //
(2) *027: e questa / la prenderei dalla parte sinistra // con due dita magari //

In the second example, the left part of the object might refer both to the handle as well as to the left part of the body of the object, but the precision grip indicated by the effector *due dita*, "two fingers", indicates that the handle is the target of the grasp. However, linguistic descriptions are sometimes insufficient to clearly understand the type of grasp (which target? which effector?) described by informants.

Therefore, for these six objects-stimuli, gestural production was also annotated, inserting an action (%act) tier in transcripts. The action tier contains a complete description of the gesture performed during the description of the grasp.[73] Only movements that can be regarded as true communicative acts are considered (for instance, not self-touching and other nonsymbolic movements). Even if the gestures transcribed have not been classified yet, we can confidently say that in this action description task most gestures produced are representational (Kita 2000; Capirci et al. 2005), in particular *enactments*, i.e. gestures in which the hands accompany speech interacting with imaginary objects and imitating motor activities; but *depictions* are also found, when the hands shape and draw the physical form or features of the real-world referent, for instance, when an informant draws a small circle in the air with a finger to represent the small handle of the cup (Kendon 2004: 160; cf. also McNeill 1992; Kita 2000). The following example reports the transcription of gestures provided from the overturned, rightward-oriented tea-cup by a right-handed informant:

*001: stesso discorso // da sotto //
*001: cioè dal fondo //
%act: solleva la mano destra all'altezza del torace e imita una presa a mano piena, con il palmo rivolto verso il basso, le dita ricurve ma piuttosto distanziate tra di loro, come attorno a un oggetto sferico; muove velocemente la mano dall'alto verso il basso due volte, poi riabbassa la mano.

[73] The kind of gesture description here adopted is similar to the one presented in Holler and Beattie (2003).

Gestures are often used to clarify verbal ambiguity (Kelly et al. 1999; Holler and Beattie 2003) and sometimes compensate the lack of information for those descriptions in which the target of the grasp is not explicitly named (i.e. 35 cases out of 180; cf. §3.2.3). In these cases, gestures express a part of the speaker's representation of the event described that speech leaves out (McNeill 1992: 79), as in the following example, provided for the leftward-oriented jug:

*028: di qua / con la sinistra //
*028: cioè / perché mi viene spontaneo farlo così //
%act: solleva la mano sinistra all'altezza del torace, stringe tutte le dita chiudendo la mano a pugno, poi riabbassa la mano.

It is clear, that even if this description contains only a reference to the effector of the grasp, the informant is still referring to a grasp directed to the handle of the object (the only part of the jug that can be grasped in a fist).

Therefore, as already stated, in this case study the 180 grasp descriptions were classified simply distinguishing between grasps that are directed to the object's handle and grasp that do not involve the handle (e.g. a grasp from the above or a lateral grasp). Moreover, when the grasp was directed to the handle, the analysis also considered whether the hand intended to grasp the handle was the right or the left one (when this kind of information is present in transcripts, in the main or in the action tier), and the dominant or non-dominant hand of the informant (§2.4.2). The purpose is to investigate the effect of object orientation and hand-dominance on the choice of the handle as the target of the grasp.

### 6.2.2 Results and discussion

The results of the classification are reported in Tab. 6.1. For the two groups of right-handed/left-handed participants, the table reports the frequency with which the handle, or another object's part, is chosen as the target of the grasp, as well as, only for the grasps directed to the handle and when this information is found in transcripts, the effector hand (right hand vs. left hand).

| Dominance | Object | Object orientation | Target = handle | | | Target = non-handle |
|---|---|---|---|---|---|---|
| | | | Tot. | Right hand | Left hand | |
| Right-handed | Tea cup | Rightward | 21 | 18 | - | 2 |
| | | Leftward | 16 | 3 | 9 | 7 |
| | Jug | Rightward | 22 | 21 | - | 1 |
| | | Leftward | 18 | 3 | 9 | 5 |
| | Ov. tea-cup | Rightward | 4 | 4 | - | 19 |
| | | Leftward | 4 | 1 | 2 | 19 |
| Left-handed | Tea cup | Rightward | 6 | 5 | - | 1 |
| | | Leftward | 7 | - | 6 | - |
| | Jug | Rightward | 7 | 4 | - | - |
| | | Leftward | 7 | - | 5 | - |
| | Ov. tea-cup | Rightward | 2 | 2 | - | 5 |
| | | Leftward | 4 | 2 | 1 | 3 |

Table 6.1. Types of grasp referred to by informants (by means of gestural and linguistic production).

It is evident that, for the jug and the upright tea-cup (120 descriptions), in both groups there is a strong, general tendency to prefer the handle as the target of the grasp (104 cases, 87%), rather than any other part of the object (16 cases, 13%). This is in line with what we have already observed in the detailed analysis of linguistic descriptions (§4.3.3.1), i.e. that the handle is the most frequent target-related word named for these objects-stimuli. However, for the overturned tea-cup, in most cases (46 cases out of 60, 77%), the grasp described by both right-handed and left-handed subjects is directed to the upper part of the object, or to the opposite side with respect to the handle, especially when the handle is not spatially aligned with their dominant hand.

Such results, again, comply with the analysis conducted in Chapter IV, which revealed that the handle was rarely mentioned as target of the grasp for the two overturned tea-cups stimuli. This confirms the relevance of habitual interaction with objects (cf. Tucker and Ellis 2004; Borghi and Riggio 2009): we usually take a tea-cup in its canonical position (i.e. upright) by its handle in order to drink something (which is the main function of a cup). However, when the object is overturned, it is often because, for instance, it has just been washed and we need to move it in another place, or just to overturn it again.

A second comment that we can make on data presented in Tab. 6.1 regards the 95 cases (over the 118 cases in which the handle is the target of the grasp described by informants)

for which we are able to identify, from linguistic descriptions and gestures, the hand that is intended to grasp the handle (Tab. 6.1, columns 5-6). Object orientation has a very strong effect on the choice of the effector hand directed to the handle, in both groups of subjects: in 86 cases out of 95, i.e. 90.5% of the descriptions, the effector hand is spatially aligned with the handle. These data confirm what the analysis of linguistic production already suggested (§4.7.1.2), that the higher frequency of explicit references to the right or the left hand, produced for the rightward-oriented or the leftward-oriented objects-stimuli respectively, is a reflex of a spatial alignment effect. This is even more interesting if we compare such results with the findings obtained in neurophysiological and behavioural studies, where a larger sensory-motor activation, or faster and more accurate responses, are observed for the hand that is spatially aligned with the affording part of the object-stimulus visually presented at subjects (§1.3.2).

With comparison to the spatial alignment, hand dominance seems to have only a slight effect on the choice of the handle as the target of the grasp: there are more grasps directed to the right-oriented handle, rather than to the left-oriented handle (47 cases vs. 38 cases), among those described by right-handed informants, and conversely, directed to the left-oriented handle, rather than to the right-oriented handle (18 cases vs. 15 cases), among those described by left-handed informants.[74]

To better visualise the results of this analysis, we can consider Figure 6.2, which represents the number of grasps directed to the handle, described by the two groups of informants for the tea-cup (represented in four different orientation conditions). For each group, the leftmost tea-cup is the stimulus for which the handle is most frequently chosen as the target of the grasp. It clearly appears that the more the handle moves away from its "canonical" position (i.e. upright tea-cup, handle oriented towards the informant's dominant hand), the more rarely it is described as the target of the action (cf. frequency data reported in Tab. 6.1):

[74] Interestingly, hand dominance seems to interfere with spatial alignment: in the analysis for the right-handed subjects, the right hand sometimes occurs in the description of a grasp directed toward leftward-oriented objects (seven cases out of 70), whereas we would have expected the left hand.


Figure 6.2. Hand dominance influences the choice to grasp the handle of the tea-cup. Each red point represents a description of a grasp directed to the handle (cf. Table 6.1).

The following table summarises the results of the analysis of the 95 grasp descriptions for which we are able to decide whether the right or left hand indicated as the effector of a grasp directed to the handle is spatially aligned with the object's orientation (±spatial alignment), also considering if it is the informant's dominant hand (±hand dominance).

| | + hand dominance | - hand dominance |
|---|---|---|
| + spatial alignment | 55 | 31 |
| - spatial alignment | 7 | 2 |

Table 6.2. Spatial alignment and hand dominance in the 95 grasp descriptions for which the transcripts provide information about which of the two hands is indicated as the effector (over the total number of 118 cases in which the handle is the target of the grasp described by informants).

Table 6.2 reveals that in this action description task, for these six stimuli, the handle is almost always chosen as the target when its orientation coincides with the dominant hand. If a conflict occurs between spatial compatibility and hand dominance, spatial compatibility strongly predominates (31 vs. 7 cases). Only two descriptions are referred to a grasp directed to a handle, performed by an effector that is neither spatially aligned with the object orientation, nor does it involve the dominant hand.

Relying on these considerations, we can conclude that the more a given type of grasp is on the left part of the following hierarchy, the more likely it is to be preferred:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| + spatial alignment<br>+ hand dominance | › | + spatial alignment<br>- hand dominance | › | - spatial alignment<br>+ hand dominance | › | - spatial alignment<br>- hand dominance |

Neurophysiological and behavioural research highlights the salience of an object's affording part with respect to other possible graspable parts, also revealing the existence of a spatial compatibility effect occurring between the orientation of the object and the ipsilateral hand, with a possible interference of hand dominance (§1.3.2). Interestingly, such phenomena appear not only to modulate motor responses, but also to influence the grasp descriptions provided by informants in a complex linguistic task, during which the memory of real interactions with objects, associated to precise motor patterns, is probably re-enacted.

## 6.3 Future directions

The research conducted in this thesis has suggested several possible applications and highlighted interesting potential future directions.

The speech corpus collected in this study from the action description task can be further enriched with additional information: in particular, it can be integrated with gesture transcription and classification. This would allow us to explore a very interesting research field, i.e. the complex interplay between syntactic, semantic, but especially prosodic aspects of speech and non-verbal behaviour (cf. for instance Jannedy and Mendoza-Denton 2006; Loehr 2012; Guellaï et al. 2014). The transcription system already adopted (cf. §2.4.5) segments the speech into utterances by identifying prosodic boundaries. Therefore, it would be interesting to investigate which kind of relation exists between the semantic content of the descriptions and the kind of gesture performed by informants, as well as whether, and to what extent, the prosodic units are aligned with gestures.

Moreover, the very small part of the corpus for which the transcription of gestures has been performed revealed that most participants often shaped the hands in the very same way, when describing the grasp of a given object-stimulus: many descriptions reported in the %act tier are almost identical to one another. In this regard, it would be of great interest to compare the gestural production of the speakers that took part in the interviews with the grasp descriptions provided for the same objects-stimuli (or just for some of them) by signing people. This comparison would surely open a window on the theme of *iconicity* (i.e. the similarity between the form of a sign and its meaning), which plays a fundamental role both

in sign languages (mainly due to the nature of their visual-gestural modality: e.g. Mandel 1977; Taub 2001; Pietrandrea 2002; Wilcox 2004) and in gestures that accompany and in most cases integrate the oral language.

Another important application scenario driven by gesture transcription is the possibility to build a grasp classification worth comparing with the extant grasp taxonomies. A first case study on gesture transcription has already been conducted on the first description provided for six objects-stimuli by all informants involved in the action description task, and it yielded interesting results. Moreover, even if this thesis was focussed only on the participant's first answer (because we were mainly interested in the most immediate type of micro-interaction afforded by objects), more than one grasp description was generally produced by each informant for each stimulus (cf. §2.4.4). For this reason, a classification based on non-verbal behaviour could not only distinguish among the different grasp types that speakers produced by gestural means, but also allow to rank specific types of grasps taking into account both the frequency and the order with which they were produced for a given stimulus or a given class of stimuli.

Such a grasp classification would have two peculiarities. The first one is related to being empirically derived from gestures performed by 30 distinct subjects, both right and left-handed, with regard to a quite numerous set of objects-stimuli (42), different from one another in terms of consistency (solid, aggregates of particles, substances), shape (objects with and without protruding parts) and size (small and large objects), affording both unimanual and bimanual grasps.

However, the strong point of this classification would probably be its association with a set of linguistic description of grasps orally produced. This permits one to align and connect two different levels:

- The level of natural language - constituted by grasp descriptions, representative of the rich way an act of grasping could be referred to with linguistic means
- The level of action execution - constituted by gestures, which, taken alone, can be seen as pantomimes of grasps

The coupling of these two levels offers a possible exploitation in a human-robot interaction scenario. The descriptions of grasps expressed in natural language could indeed be modelled as, and converted into, a set of "instructions", a base of knowledge related to

how to grasp objects in a human-like manner, whereas the gesture produced by informants in association with such instructions could be mapped onto bio-inspired models of grasping adopted for the design and the control of robotic hands. Such models aim at achieving simpler and more performing system architectures for the hand, through which robot hands may be capable of adapting more robustly to different tasks and different environments.[75] This field, which nowadays attracts a broad community effort from neurosciences, cognitive sciences, robotics, engineering and physics, represents a fertile soil for the affordance theory (cf. also §1.2.2).

These are only some of the possible applications suggested by the research conducted that might be possibly explored in collaboration with other research units of the ModelAct project: in particular, with the group of the Institute of Cognitive Sciences and Technologies (ISTC, CNR, Rome), especially for what regards gestures and sign language, and with the Siena Robotics and Systems Lab (SIRSLab) and the BioRobotics Institute of Scuola Superiore Sant'Anna (Pontedera, Pisa), regarding possible applications in research on bio-inspired grasp models.

[75] See for instance the research recently conducted within the European projects *THE Hand Embodied* (http://www.thehandembodied.eu/).

# Conclusion

The main results of this work have already been presented and discussed in the final sections of the previous chapters. However, since many different issues were tackled, it will be useful to briefly retrace the path of this research and to highlight its main contribution.

In Chapter I, we reviewed a body of neurophysiological and behavioural studies that provide convincing evidence for the existence of a close relationship between perceptual and motor processes. In particular, a number of experiments demonstrate that the simple visual perception of a graspable object is able to automatically evoke a sort of "action simulation" in the motor system, activating the same neural circuits that fire during object manipulation and grasping (e.g. Grafton et al. 1997; Binkofski et al. 1999; Chao and Martin 2000; Grèzes et al. 2003; Buccino et al. 2009). Since the early involvement of the motor system after the visual presentation of a stimulus is able to influence the performance of subsequent actions (causing interference or facilitation effects), action simulations may be considered as affordances, intended as true «preconditions for action» (Greeno 1994: 340). Following from these empirical findings, we defined affordances as the motor simulations of possible actions incorporated in the objects' representations that are automatically triggered by the perception of visually presented stimuli (cf. Ellis and Tucker 2000; Grèzes, Tucker et al. 2003).

However, this recruitment of the sensory-motor system is modulated by a number of factors, strictly related to agents' real experiences and past interactions with objects. Artefacts

and tools are typically involved in the actions of manipulation and grasping and, when used as experimental stimuli, they are more effective in activating motor simulations with respect, for example, to natural kinds. Moreover, within the artefacts category, there are specific object features that are able to influence sensory-motor responses. In particular, in order to use a tool, agents generally have to manipulate and grasp its affording part (e.g. its handle). Therefore, as a consequence of repeated interactions with the object, the presence of an affording part usually causes a greater activation in the sensory-motor regions, especially when such object's component is spatially aligned with the hand for which brain activity is measured (e.g. Grèzes et al. 2003; Buccino et al. 2009; cf. also Tucker and Ellis 1998; Phillips and Ward 2002; Symes et al. 2007).

In Chapter II, an action description task explicitly designed to investigate linguistic reflexes of the affordance of graspability has been illustrated. During the interviews, subjects were visually presented with 42 pictures of graspable entities and were asked to describe how they would grasp such entities. The results of this experiment, discussed in Chapters III and IV, well comply with the findings from psychological and neuroscientific research presented in the first chapter. The target of the grasp is named in most grasp descriptions provided for artefacts and humans, whereas it is less frequently mentioned for substances and natural objects. Moreover, within the category of artefacts, the target is explicitly indicated more often for artefacts provided with affording parts, with respect to artefacts without affording parts. Therefore, we have described the distribution of the references to the target and the effector of the grasp with a hierarchy, according to which the more an object-stimulus is on the left part of this hierarchy, the more likely the target of the grasp is to be named in the descriptions, and the less likely the effector of the grasp is to be mentioned: *human kinds* > *artefacts with AP* > *artefacts without AP* > *natural kinds* > *substances/aggregates*. Moreover, even if the target of the grasp is generally indicated by means of lexical expressions pertaining to the spatial domain, we have observed that for the categories of human kinds and artefacts with affording parts participants usually produced meronyms (which in most cases denote objects' affording parts).

In Chapter V, we have compared the meronyms produced in this action description task with those listed in a property generation task focussed on the category of artefacts with affording parts. The stimuli consisted of written words and informants were asked to list a series of features that they considered relevant in order to describe the meaning of the linguistic items. Results reveal that, on the one hand, meronyms denoting the object's parts

most typically involved in grasp and manipulation (i.e. handles) are the most frequently produced, both in the action description and in the property generation tasks. However, affording parts explicitly named in the first task are much more numerous: since the experiment was specifically focussed on the action of grasping, participants' attention was drawn to the affording parts purposely designed for grasping and clearly related to this specific function.

Finally, in Chapter VI, we have presented some possible applications of the results of the research and commented on possible future directions. In particular, we have proposed that information related to affordances could prove useful within the field of computational linguistics and tested the validity of this hypothesis annotating the information related to affordances in a corpus, in order to help disambiguating between concrete and metaphorical uses of verbs of grasping (e.g. *to grasp a pen* vs. *to grasp an idea*). A second case study explored the possibility to enrich the corpus collected from the interviews with additional information and proved the utility of gesture annotation to determine the effects of object's orientation and hand dominance on grasp descriptions.

In conclusion, the main purpose of this research was to investigate whether the language reflects affordances. The analysis conducted on the grasp descriptions has revealed that linguistic behaviour results to be influenced by the same factors that behavioural and neurophysiological researches indicate as able to modulate motor responses, namely the objects' semantic type and the salience of the affording part with respect to other possible graspable parts of the objects. Moreover, also the spatial compatibility effect occurring between the orientation of the object and the informant's ipsilateral hand, together with hand dominance, has an effect of the linguistic description of grasp, that become even more evident when taking into consideration also gesture production. Describing an action requires an imagery process, during which the experience of concrete interactions with objects is re-enacted; the same happens when action simulations are automatically triggered by object perception even if no actual reach-and-grasp movement is executed.

Since its outset, this thesis has been characterised by a multidisciplinary approach to the theme of language and affordances. This is partly due to the multifaceted concept of affordance itself, which may function as a bridge that crosses the boundaries of different disciplines. After its first appearance within the ecological psychology framework, this notion soon made huge inroads into other fields; however, affordances have been extensively

investigated in psychology and neuropsychology, whereas their potential is still largely underestimated in linguistics.

If, on the one hand, recent trends in neurosciences and psychology have provided new stimuli to linguistic studies, on the other hand, linguists can make an important contribution on issues that, so far, have predominantly been explored with different approaches. In this regard, we hope that this work may contribute to drawing attention not only to how the notion of affordance could be fruitfully applied to the linguistic domain, but also to how a linguistic perspective on this theme can shed new light on the findings that have emerged from other research fields.

# Appendix A

The pictures used as stimuli in the action description task (cf. §2.4.3.2).







# Appendix B

Translation of the instructions provided to informants in the property generation task (cf. §5.2.2).

Thank you for agreeing to participate in this survey!
In this survey, you will be presented with 14 words that indicate common use objects, each one in a different page. Below each word, you will see some blank lines. Here, you will have to describe the meaning of the word, using up to a maximum of 10 simple short sentences, as shown in the following example.

DOG:
- is a mammal;
- is a man's best friend;
- barks;
- is a pet;
- has a tail.

RULES: The task you are required is very simple, I am sure it will not take you a lot of time. However, you have to complete it accurately, so here there are some rules to follow:
1. Do not hurry. For each word, first think carefully to its meaning and to the aspects you consider most important to describe it; then, fill in the fields at your disposal.
2. Describe the meaning of the word with short sentences. Try to be clear and concise.
3. There are no right or wrong answers: you are absolutely free to explain as you prefer what in your opinion is the meaning of these words.
4. You are not obliged to fill all the 10 lines for each word.
5. Once you have completed a page, make sure that the data you entered are correct, because you cannot change them later. When you are sure, click on the "Next" button, which brings you to the next page.
6. You cannot interrupt the survey and save your answers: the survey can be saved only at the end, after its completion.

# Acknowledgments

I wish to express my sincere gratitude to my supervisors, without whom this thesis would not have been possible. To Prof. Giovanna Marotta, for the enthusiasm with which she supervised all stages of this research, for all her brilliant contributions, for always being available to clarify my doubts and to discuss ideas, and for her continuous support and guidance all along. To Dr. Nicoletta Calzolari, for providing constructive comments on this thesis, for her expert advice, and for all the conversations we had: they have always been so stimulating and helpful for me. To Dr. Monica Monachini, for her constant encouragement and appreciation, for the time she dedicated to me, for all her precious suggestions, but especially for addressing me towards the theme of affordances.

I owe my gratitude to the Department of Linguistics of the University of Pisa that gave me the opportunity to accomplish my Ph.D., and that provided facilities, tools and support for conducting the experiments.

I am grateful to the Institute for Computational Linguistics "Antonio Zampolli" (CNR, Pisa), headed by Dr. Simonetta Montemagni, not only for supporting my research activity with the fundings obtained for the ImagAct and ModelAct projects, but also for providing me the chance to work in a very stimulating environment. In particular, I have to thank Dr. Irene Russo, for her helpful comments on many parts of this thesis; Dr. Fahad Khan, for the patience in answering all my questions; Dr. Francesca Frontini, for all her advice; and all the

other people of the Resources and Linguistic Technologies Research Unit, from whom I learnt so much. I wish to thank also the other groups involved in the ModelAct project, in particular the LABLITA research group, especially Prof. Massimo Moneglia and Dr. Alessandro Panunzi, for suggestions about the transcription of the interviews, as well as the ISTC group, especially Dr. Elena Tomasuolo and Dr. Virginia Volterra, for interesting discussions on gestures and sign language.

I would like to thank the coordinator of the Ph.D. program in Linguistics of the University of Pisa, Prof. Alessandro Lenci, for helping me with the property generation task and for his valuable suggestions.

Thanks also to Dr. Gianluca Lebani, for his time and skilful advice on semantic norms; to Dr. Gloria Gagliardi, for useful comments on Chapter III; to Philippa Evans, for her careful proofreading; and to all students who voluntarily took part in the experiments.

I cannot thank my parents enough for their invaluable loving and encouraging support throughout these years, and especially during the final stages of this Ph.D. Without their emotional backing, this thesis would never have seen the light. A special thanks to my mother, my first, most careful and tireless reader.

Lastly, I wish to thank my partner, for his patience and his unfailing and unconditional love.

# References

Albano Leoni, F. 2003. Tre progetti per l'italiano parlato. In N. Maraschio, and T. Poggi Salani (eds.), *Italia linguistica anno Mille. Italia linguistica anno Duemila. Atti del XXXIV congresso internazionale di studi della Società di Linguistica Italiana (SLI)*, 675-683. Roma: Bulzoni.

Alexiadou, A., and Schäfer, F. 2006. Instrument subjects are agents or causers. In D. Baumer, D. Montero, and M. Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 40-48. Somerville: Cascadilla Proceedings Project.

Attardo, S. 2005. The role of affordances at the semantics/pragmatics boundary. In B. G. Bara, L. Barsalou, and M. Bucciarelli (eds.), *Proceedings of the XXVII annual conference of the Cognitive Science Society (CogSci 2005)*, 169-174. Mahwah: Lawrence Erlbaum Associates.

Austin, J. L. 1962. *How to do things with words*. Oxford: Oxford University Press.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology* 16: 1818-1823.

Barner, D., and Snedeker, J. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 97: 41-66.

Baroni, M., Lenci, A., and Cazzolli, G. 2013. Norme semantiche per lo studio del significato. In G. Marotta, L. Meini, and M. Donati (eds.), *Parlare senza vedere. Rappresentazioni semantiche nei non vedenti*, 67-81. Pisa: ETS.

Barsalou, L. W. 1999. Perceptual symbol systems. *Bahavioral and Brain Sciences* 22: 577: 609.

Barsalou, L. W. 2008. Grounded cognition. *Annual Review of Psychology* 59: 617-645.

Barsalou, L. W., Santos, A., Simmons, W. K., and Wilson, C. D. 2008. Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, and A. C. Graesser (eds.), *Symbols, embodiment, and meaning*, 245-283. Oxford: Oxford University Press.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 7: 84-91.

Barsalou, L. W., and Wiemer-Hastings, K. 2005. Situating abstract concepts. In D. Pecher, and R. Zwaan (eds.), *Grounding cognition: The role of perception and action in memory, language, and thought*, 129-163. New York: Cambridge University Press.

Bartolini, R., Quochi, V., De Felice, I., Russo, I., and Monachini, M. 2014. From synsets to videos: enriching ItalWordNet multimodally. In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis (eds.), *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*, 3110-3117.

Becvar, L. A., Hollan, J., and Hutchins, E. 2005. Representational gestures as cognitive artifacts for developing theory in scientific laboratory. *Semiotica* 156: 89-112.

Binkofski, F., Buccino, G., Posse, S., Seitz, R., Rizzolatti, G., and Freund, H. 1999. A fronto-parietal circuit for object manipulation in man: Evidence from an fMRI-study. *European Journal of Neuroscience* 11: 3276-3286.

Bloom, P. 1994. Intention, history, and artifact concepts. *Cognition* 60: 1-29.

Borghi, A. M. 2004. Object concepts and action: Extracting affordances from objects parts. *Acta Psychologica* 115: 69-96.

Borghi, A. M. 2005. Object concepts and action. In D. Pecher, and R. A. Zwaan (eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking*, 8-34. Cambridge: Cambridge University Press.

Borghi, A. M. 2007. Object concepts and embodiment: Why sensorimotor and cognitive processes cannot be separated. *La Nuova Critica* 15: 447-472.

Borghi, A. M. 2012. Language comprehension: Action, affordances and goals. In Y. Coello, and A. Bartolo (eds.), *Language and action in cognitive neuroscience*, 125-144. Hove/New York: Psychology Press.

Borghi, A. M., Flumini, A., Natraj, N., and Wheaton, L. A. 2012. One hand, two objects: Emergence of affordance in contexts. *Brain and Cognition* 80: 64-73.

Borghi, A. M., Glenberg, A. M., and Kaschak, M. P. 2004. Putting words in perspective. *Memory and Cognition* 32(6): 863-873.

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

Bosque-Gil, J., and Pustejovsky, J. 2015. The semantics of image annotation. In H. Bunt (ed.), *Proceedings 11th Joint ACL - ISO Workshop on Interoperable Semantic Annotation (isa-11)*, 29-38.

Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., and Nazir, T. A. 2008. Word processing in Parkinson's desease is impaired for action verbs but not for concrete nouns. *Neuropsychologia* 46: 743:756.

Bressem, J., Ladewig, S., and Müller, C. 2013. Linguistic Annotation System for Gestures (LASG). In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and S. Teßendorf (eds.), *Body - Language - Communication: An international handbook on multimodality in human interaction*, 1098-1124. Berlin/Boston: Mouton De Gruyter.

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

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and 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., and Mezzadri, M. 2013. La teoria dell'embodiment e il processo di apprendimento e insegnamento di una lingua. *Enthymema* 8: 5-20.

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

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., and Riggio, L. 2009. Broken affordances, broken objects: A TMS study. *Neuropsychologia* 47: 3074-3078.

Burgess, C. 1998. From simple associations to the building blocks of language: Modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, and Computers* 30: 188-198.

Çakmak, M. 2007. *Robot planning based on learned affordances*. Master's thesis, Middle East Technical University, Ankara. Retrieved from http://www.cc.gatech.edu/~maya/papers/cakmak_msc_thesis.pdf (last accessed: May 3, 2015).

Capirci, O., Contaldo, A., Caselli, M. C., and Volterra, V. 2005. From action to language through gesture: A longitudinal perspective. *Gesture* 5(1-2): 155-177.

Caramazza, A., Mahon, B. Z. 2003. The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends in Cognitive Sciences* 7: 354-361.

Cardellicchio, P., Sinigaglia, C., and Costantini, M. 2011. The space of affordances: A TMS study. *Neuropsychologia* 49: 1369:1372.

Cardellicchio, P., Sinigaglia, C., and Costantini, M. 2013. Grasping affordances with the other's hand: A TMS study. *Social Cognitive and Affective Neuroscience* 8(4): 455-459.

Castner, J. E., Chenery, H. J., Silburn, P. A., Coyne, T. J., Sinclair, F., Smith, E. R., and Copland, D. A. 2008. Effects of subthalamic deep brain stimulation on noun/verb generation and selection from competing alternatives in Parkinson's desease. *Journal of Neurology, Neurosurgery, and Pychiatry* 79: 700-705.

Cattaneo, Z., Devlin, J. T., Salvini, F., Vecchi, T., and Silvanto, J. 2010. The causal role of category-specific neuronal representations in the left ventral premotor cortex (PMv) in semantic processing. *NeuroImage* 49: 2728-2734.

Chao, L. L., Haxby, J. V., and Martin, A. 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience* 2: 913-919.

Chao, L. L., and Martin, A. 2000. Representation of manipulable man-made objects in the dorsal stream. *NeuroImage* 12: 478-484.

Chatterjee, A. 2010. Disembodying cognition. *Language and Cognition* 2: 79-116.

Chemero, A. 2001. What we perceive when we perceive affordances: Commentary on Michaels (2000) "Information, Perception and Action". *Ecological Psychology* 13(2): 111-116.

Chemero, A. 2003. An outline of a theory of affordances. *Ecological Psychology* 15(2): 181-195.

Chemero, A. 2009. *Radical embodied cognitive science.* Cambridge, MA: MIT press.

Chemero, A., Klein, C., and Cordeiro, W. 2003. Events as changes in the layout of affordances. *Ecological Psychology* 15(1): 19-28.

Chierchia, G. 1998. Plurality of mass nouns and the notion of "semantic parameter". In S. Rothstein (ed.), *Events and grammar,* 53-103. Dordrecht: Kluwer.

Ciaramita, M., and Altun, Y. 2006. Broad-coverage sense disambiguation and information extraction with a supersense sequence tagger. In D. Jurafsky, and E. Gaussier (eds.), *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, 594-602. Association for Computational Linguistics.

Ciaramita, M., and Johnson, M. 2003. Supersense tagging of unknown nouns in WordNet. In M. Collins, and M. Steedman (eds.), *Proceedings of the 2003 Empirical Methods in Natural Language Processing (EMNLP 2003)*, 168-175. Association for Computational Linguistics.

Clausen, D., Djalali, A., Grimm, S., Lauer, S., Rojas-Esponda, T., and Levin, B. 2010. *Extension, ontological type, and morphosyntactic class: Three ingredients of countability*. Extended abstract. Stanford University. Retrieved from http://web.stanford.edu/~bclevin/bochum10abst.pdf (last accessed: May 3, 2015).

Cocks, N., Dipper, L., Pritchard, M., and Morgan, G. 2013. The impact of impaired semantic knowledge on spontaneous iconic gesture production. *Aphasiology* 27(9): 1050-1069.

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., and Committeri, G. 2010. Where does an object trigger an action? An investigation about affordances in space. *Experimental Brain Research* 207: 95-103.

Costantini, M, Committeri G, and Sinigaglia, C. 2011. Ready both to your and to my hands: Mapping the action space of others. *PLoS ONE* 6(4): e17923. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0017923 (last accessed: May 3, 2015).

Costantini, M., and Sinigaglia, C. 2012. Grasping affordance: A window onto social cognition. In A. Seeman (ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*, 431-470. Cambridge, MA: MIT Press.

Coyne, B., Klapheke, A., Rouhizadeh, M., Sproat, R., and Bauer, D. 2012. Annotation tools and knowledge representation for a text-to-scene system. In *Proceedings of COLING 2012: technical papers*, 679-694. The COLING 2012 Organizing Committee.

Coyne, B., and Sproat R. 2001. WordsEye: An automatic text-to-scene conversion system. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 487-496.

Cree, G., and McRae, K. 2003. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology. General* 132(2): 163-201.

Creem, S. H., and Proffitt, D. R. 2001. Grasping objects by their handles: A necessary interaction between cognition and action. *Journal of Experimental Psychology: Human Perception and Performance* 27(1): 218-228.

Crescentini, C., Mondolo, F., Biasutti, E., and Shallice, T. 2008. Supervisory and routine processes in noun and verb generation in nondemented patients with Parkinson's desease. *Neuropsychologia* 46: 434-447.

Cresti, E. 1994. Information and intonation patterning. In P. Martin, B. G. Ferguson, and H. Gezundhajt (eds.), *Accent, intonation and modèles phonologiques*, 99-140. Toronto: Edition Mélodie.

Cresti, E. 2000. *Corpus di italiano parlato*. Firenze: Accademia della Crusca.

Cresti, E., and Moneglia, M. (eds.). 2005. *C-ORAL-ROM. Integrated reference corpora for spoken Romance languages*. Amsterdam: John Benjamins.

Cruse, D. A. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Crystal, D. 1975. *The English tone of voice*. London: Edward Arnold.

Cutkosky, M. R., and Wright, P. K. 1989. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation* 5(3): 269-279.

De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., and Storms, G. 2008. Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods* 40(4): 1030-1048.

De Felice, I. 2013. Affordances: Una chiave per il word sense disambiguation. In A. Auricchio, M. Cruciani, A. Rega, and M. Villani (eds.), *Special issue: Atti del X convegno annuale dell'Associazione Italiana di Scienze Cognitive. Nea Science* 1(2): 105-110.

De Felice, I. 2014a. From hands to handles: How objects' orientation affects grasp descriptions. In A. Auricchio, M. Cruciani, A. Rega (eds.), *Special issue: atti del XI convegno annuale dell'Associazione Italiana di Scienze Cognitive. Nea Science* 5: 119-125.

De Felice, I. 2014b. "Possibilities of action" in language: Affordances and verbal polysemy. *Reti, Saperi, Linguaggi. Italian Journal of Cognitive Sciences* 1: 179-191.

De Felice, I., Bartolini, R., Russo, I., Quochi, V., and Monachini, M. 2014. Evaluating ImagAct-WordNet mapping for English and Italian through videos. In R. Basili, A. Lenci, and B. Magnini (eds.), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and the Fourth International Workshop EVALITA 2014*, 128-131. Pisa: Pisa University Press.

De Mauro, T., Mancini, F., Vedovelli, M., and Voghera, M. 1993. *Lessico di frequenza dell'Italiano parlato*. Milano: ETASLIBRI.

De Vega, M. 2012. Language and action: An approach to embodied cognition. In V. Gyselinck, and F. Pazzaglia (eds.), *From mental imagery to spatial cognition and language*, 177-199. Hove/New York: Psychology Press.

Di Pellegrino, G., Fadiga, L., Fogssi, L., Gallese, V., and Rizzolatti, G. 1992. Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91(1): 176-180.

Dressler, W. U., and Barbaresi, L. M. 1994. *Morphopragmatics: Diminutives and intensifiers in Italian, German, and other languages*. Berlin: Mouton de Gruyter.

Duncan, S. 2005. Co-expressivity of speech and gesture: Manner of motion in Spanish, English, and Chinese. In C. Chang, M. J. Houser, Y. Kim, D. Mortensen, M. Park-Doob,

and M. Toosarvandani (eds.), *Proceedings of the 27th Annual Meeting of the Berkeley Linguistic Society*, 353-370. Berkeley: Berkeley Linguistics Society.

Ellis, R., and Tucker, M. 2000. Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology* 91: 451-471.

Enticott, P., Kennedy, H., Bradshaw, J., Rinehart, N., and Fitzgerand, P. 2010. Understanding mirror neurons: Evidence for enhanced corticospinal excitability during the observation of transitive but not intransitive hand gestures. *Neuropsychologia* 48: 2675-2680.

Fadiga, L., and Craighero, L. 2003. New insights on sensorimotor integration: From hand action to speech perception. *Brain and Cognition* 53: 514-524.

Fagg, A. H., and Arbib, M. A. 1998. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 11: 1277-1303.

Feix, T., Pawlik, R., Schmiedmayer, H., Romero, J., and Kragic, D. 2009. A comprehensive grasp taxonomy. In *Proceedings of the Robotics, Science and Systems Conference. Workshop on Understanding the Human Hand for Advancing Robotic Manipulation.* Retrieved from http://www.csc.kth.se/~danik/taxonomyGRASP.pdf (last accessed: May 3, 2015).

Fellbaum, C. 1998. *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Field, A. 2013[4]. *Discovering statistics using SPSS*. London: Sage Publications.

Fischer, M. H., and Zwaan, R. A. 2008. Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology* 61(6): 825-50.

Fodor, J. A. 1975. *The language of thought*. Cambridge: Harvard University Press.

Frassinelli, D., and Lenci, A. 2012. Concepts in context: Evidence from a feature-norming study. In N. Miyake, D. Peebles, and R. P. Cooper (eds.), *Building bridges across cognitive sciences around the world. Proceedings of the 34th annual meeting of the Cognitive Science Society (CogSci)*, 1567-1571. Austin: Cognitive Science Society.

Frontini, F., De Felice, I., Russo, I., Khan, F., Monachini, M., Gagliardi, G., and Panunzi, A. 2012. Verb interpretation for basic action types: Annotation, ontology induction and

creation of prototypical scenes. In M. Zock, and R. Rapp, *Proceedings of the 3rd Workshop on Cognitive Aspects of the Lexicon (CogALex III)*, 69-80. Curran Associates.

Gagliardi, G. 2014. *Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnosi del Mild Cognitive Impairment (MCI)*. Doctoral dissertation, University of Florence, Florence. Retrieved from http://www.gloriagagliardi.com/wp-content/uploads/2013/12/Tesi_Gloria_color.pdf (last accessed: May 3, 2015).

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

Gaver, W. W. 1991. Technology affordances. In S. P. Robertson, G. M. Olson, and J. S. Olson (eds.), *Proceedings of the CHI '91 human factors in computing systems conference*, 79-84. New Orleans: ACM Press.

Gentilucci, M. 2002. Object motor representation and reaching-grasping control. *Neuropsychologia* 40: 1139-1153.

Gibbs, R. W. 2003. Embodied experience and linguistic meaning. *Brain and Language* 84: 1-15.

Gibson, E. J. 2000a. Where is the information for affordances? *Ecological Psychology* 12(1): 53-57.

Gibson, E. J. 2000b. Perceptual learning in development: Some basic concepts. *Ecological Psychology* 12(4): 295-302.

Gibson, E. J. 2003. The world is so full of a number of things: On specification and perceptual learning. *Ecological Psychology* 15(4): 283-287.

Gibson, E. J., and Pick, A. 2000. *An ecological approach to perceptual learning and development*. Oxford: Oxford University Press.

Gibson, J. J. 1966. *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. 1977. The theory of affordances. In R. Shaw, and J. Bransford (eds.), *Perceiving, acting, and knowing. Toward an ecological psychology*, 67-82. Hillsdale: Lawrence Erlbaum Associates.

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

Girardi, G., Lindemann, O., and Bekkering, H. 2010. Context effects on the processing of action-relevant object features. *Journal of Experimental Psychology* 36: 330-340.

Glenberg, A. M. 2010. Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science* 1: 586-596.

Glenberg, A. M., and Kaschak, M. P. 2002. Grounding language in action. *Psychonomic Bullettin and Review* 9: 558-565.

Glover, S., Rosenbaum, D. A., Graham, J., and Dixon, P. 2004. Grasping the meaning of words. *Experimental Brain Research* 154: 103-108.

Gorniak, P., and Roy, D. 2006. Perceived affordances as a substrate for linguistic concepts. In R. Sun (ed.), *Proceedings of the 29th annual cognitive science society*, 279-284.

Gorniak, P., and Roy, D. 2007. Situated language understanding as filtering perceived affordances. *Cognitive Science* 31(2): 197-231.

Gough, P. M., Riggio, L., Chersi, F., Sato, M., Fogassi, L., and Buccino, G. 2012. Nouns referring to tools and natural objects differentially modulate the motor system. *Neuropsychologia* 50: 19-25.

Grabowski, T. J., Damasio, H., and Damasio, A. R. 1998. Premotor and prefrontal correlates of category-related lexical retrieval. *NeuroImage* 7: 232-243.

Graffi, G. 1994. *Sintassi*. Bologna: Il Mulino.

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

Greeno, J. 1994. Gibson's affordances. *Psychological Review* 101(2): 336-342.

Grèzes, J., Armony, J., Rowe, J., and Passingham, R. 2003. Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *NeuroImage* 18: 928-937.

Grèzes, J., and Decety, J. 2002. Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* 40: 212-222.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., and Passingham, R. 2003. Objects automatically potentiate action: An fMRI study of implicit processing. *European Journal of Neuroscience* 17: 2735-2740.

Guellaï, B., Langus, A., and Nespor, M. 2014. Prosody in the hands of the speaker. *Frontiers in Psychology* 5: 700.

Gullberg, M., De Bot, K., and Volterra, V. 2008. Gestures and some key issues in the study of language development. *Gesture* 8(2): 149-179.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. 2009. The WEKA data mining software: An update. *SIGKDD Explorations* 11(1): 10-18.

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

Havas, D. A., Glenberg, A. M., Gutowsky, K. A., Lucarelli, M. J., and Davidson, R. J. 2010. Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science* 21(7): 895-900.

Havas, D. A., Glenberg, A. M., and Rinck, M. 2007. Emotion simulation during language comprehension. *Psychonomic Bulletin and Review* 14: 436-441.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425-2430.

Holler, J., and Beattie, G. 2003. Pragmatic aspects or representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture* 3(2): 127-154.

Horton, T. E., Chakraborty, A., and St. Amant, R. 2012. Affordances for robots: A brief survey. *Avant* 3(2): 70-84.

Humphreys, G. 2001. Objects, affordances... action! *The psychologist* 14(8): 408-412.

Jackendoff, R. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jannedy, S., and Mendoza-Denton, N. 2006. Structuring information through gesture and intonation. *Interdisciplinary Studies on Information Structure* 3: 199-244.

Jeannerod, M. 1994. The representing brain: Neural correlates of motor intention and imagery. *Behavioural Brain Science* 17: 187-245.

Jeannerod, M., Arbib, M., Rizzolatti, G., and Sakata, H. 1995. Grasping objects: The cortical mechanisms of visuomotor transformation. *Trends in Neurosciences* 18(7): 314-320.

Ježek, E., and Varvara, R. 2015. Instrument subjects without Instrument role. In H. Bunt (ed.), *Proceedings of the 11th Joint ACL - ISO Workshop on Interoperable Semantic Annotation (isa-11)*, 46-54.

Ježek, E., Vieu, L, Zanzotto, F. M., Vetere, G., Oltramari, A., Gangemi, A., and Varvara, R. 2014. Enriching Senso Comune with Semantic Role Sets. In H. Bunt (ed.), *Proceedings of the 10th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (isa-10)*, 88-94.

Jirak, D. Menz, M. M., Buccino, G., Borghi, A. M., and Binkofski, F. 2010. Grasping language: A short story on embodiment. *Consciousness and Cognition* 19: 711-720.

Kang, S. B., and Ikeuchi, K. 1992. Grasp recognition using the contact web. In *Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 194-201.

Karcevsky, S. 1931. Sur la phonologie de la phrase. *Travaux du Cercle linguistique de Prague* 4: 188-228.

Kaschak, M., and Glenberg, A. 2000. Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language* 43: 508-529.

Kelly, S. D., Barr, D. J., Church, R. B., and Lynch, K. 1999. Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language* 40: 577-592.

Kemmerer, D. 2006. Action verbs, argument structure constructions, and the mirror neurons system. In A. Arbib (ed.), *Action to language via the mirror neuron system*, 347-373. Cambridge: Cambridge University Press.

Kendon, A. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., and Hare, R. D. 1999. Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping* 7: 225-233.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., and Suchomel, V. 2014. The Sketch Engine: Ten years on. *Lexicography* 1: 7-36.

Kipp, M. 2004. *Gesture generation by imitation: From human behavior to computer character animation*. Boca Raton: Dissertation.com.

Kita, S. 2000. How representational gestures help speaking. In D. McNeill (ed.), *Language and gesture*, 379-415. New York: Cambridge University Press.

Klima, E. S. 1965. *Studies in diachronic syntax*. Doctoral dissertation, Harvard University, Harvard.

Koffka, K. 1935. *Principles of gestalt psychology*. New York: Wiley.

Kremer, G., and Baroni, M. 2011. A set of semantic norms for German and Italian. *Behavior Research Methods* 43(1): 97-119.

Lebani, G. E., and Pianta, E. 2010. A feature type classification for therapeutic purposes: A preliminary evaluation with non-expert speakers. In *Proceedings of the Fourth Linguistic Annotation Workshop*, 157-161. Association for Computational Linguistics.

LeBaron, C., and Streeck, J. 2000. Gestures, knowledge and the world. In D. McNeill (ed.), *Language and gesture*, 118-138. New York: Cambridge University Press.

Lenci, A., Baroni, M., Cazzolli, G., and Marotta, G. 2013. BLIND: A set of semantic feature norms from the congenitally blind. *Behavior Research Methods* 45(4): 1218-1233.

Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowsky, A., Peters, I., Peters, W., Ruimy, N., Villegas, M., and Zampolli, A. 2000. SIMPLE: A general framework for the development of multilingual lexicon. *International Journal of Lexicography* 13(4): 249-263.

Levin, B. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Lindsey, R., Veksler, V. D., Grintsvayg, A., and Gray, W. D. 2007. Be wary of what your computer reads: The effects of corpus selection on measuring semantic relatedness. In *Proceedings of the 8th International Conference on Cognitive Modeling*, 279-284.

Linkenauger, S. A., Witt, J. K., Stefanucci, J. K., Bakdash, J. Z., and Proffitt, D. R. 2009. The effect of handedness and reachability on perceived distance. *Journal of Experimental Psychology: Human Perception and Performance* 35: 1649: 1660.

Liu, J., Feng, F., Nakamura, Y. C., and Pollard, N. S. 2014. A taxonomy of everyday grasps in action. In *Proceedings of the 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2014),* 573-580.

Liuzza, M. T., Setti, A., and Borghi, A. M. 2012. Kids observing other kids' hands: Visuomotor priming in children. *Consciousness and Cognition* 21: 383-392.

Loehr, D. P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology* 3(1): 71-89.

Mace, W. M. 1977. James J. Gibson's strategy for perceiving: Ask not what's inside your head, but what your head's inside of. In R. Shaw, and J. Bransford (eds.), *Perceiving, acting, and knowing. Toward an ecological psychology*, 43-65. Hillsdale: Lawrence Erlbaum Associates.

MacKenzie, C. L., and Iberall, T. 1994. *The grasping hand*. Amsterdam: North-Holland.

MacWhinney, B. 1999. The emergence of language from embodiment. In B. MacWhinney (ed.), *The emergence of language*, 213-256. Mahwah: Lawrence Erlbaum Associates.

MacWhinney, B. 2000[3]. *The CHILDES Project: Tools for analyzing talk*. Mahwah: Lawrence Erlbaum Associates.

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

Mandel, M. 1977. Iconic devices in American Sign Language. In L. Friedman (ed.), *On the other hand: New perspectives on American Sign Language*, 57-107. New York: Academic Press.

Marino, B., Gough, P., Gallese, V., Riggio, L., and Buccino, G. 2013. How the motor system handles nouns: A behavioural study. *Psychological Research* 77: 64-73.

Mark, L. S. 1987. Eye height scaled information about affordances: A study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 13: 361-370.

Mark, L. S., Balliett, J., Craver, K., Douglas, S. D., and Fox, T. 1990. What an actor must do to perceive the affordance for sitting. *Ecological Psychology* 2: 325-366.

Mark, L. S., Jiang, Y., King, S. S., and Paasche, J. 1999. The impact of visual exploration on judgments of whether a gap is crossable. *Journal of Experimental Psychology: Human Perception and Performance* 25: 287-295.

Marotta, G. 2013. Linguaggio, cognizione e visione. In G. Marotta et al. (eds.), *Parlare senza vedere. Rappresentazioni semantiche nei non vedenti*, 13-34. Pisa: ETS.

Marotta, G., Meini, L., and Donati, M. (eds.). 2013. *Parlare senza vedere. Rappresentazioni semantiche nei non vedenti*. Pisa: ETS.

Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. 1996. Neural correlates of category-specific knowledge. *Nature* 379: 649-652.

Martín-Loeches, M., Hinojosa, J. A., Fernández-Frías, C., and Rubia, F. J. 2001. Functional differences in the semantic processing of concrete and abstract words. *Neuropsychologia* 39: 1086-1096.

McDonald, D., and Pustejovsky, J. 2014. Representing inferences and their lexicalization. *Advances in Cognitive Systems* 3: 143-162.

McEnery, A. M., and Wilson, A. 2001[2]. *Corpus linguistics*. Edinburgh: Edinburgh University Press.

McGrenere, J., and Ho, W. 2000. Affordances: Clarifying and evolving a concept. In *Proceedings of Graphics Interface 2000*, 179-186.

McNally, L. and Boleda, G. (submitted). Conceptual vs. referential affordance in concept composition. In Y. Winter, and J. Hampton (eds.), *Concept composition and experimental semantics/pragmatics*. Springer, Berlin. Retrieved from http://www.upf.edu/pdi/louise-mcnally/_pdf/publications/McNally-Boleda-14.pdf (last accessed: May 3, 2015).

McNeill, D. 1992. *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McRae, K., Cree, G. S., Seidenberg, M. S., and McNorgan, C. 2005. Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods* 37(4): 547-559.

Meini, L. 2010. *Dimensioni dello spazio nelle preposizioni. Uno studio empirico sull'italiano L2*. Pisa: PLUS - Pisa University Press.

Michaels, C. 2003. Affordances: Four points of debate. *Ecological Psychology* 15: 135-148.

Middleton, E. L., Wisniewski, E. J., Trindel, K. A., and Imai, M. 2004. Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language* 50: 371-394.

Moneglia, M. 2005. The C-ORAL-ROM resource. In E. Cresti, and M. Moneglia (eds.), *C-ORAL-ROM. Integrated reference corpora for spoken Romance languages*, 1-70. Amsterdam: John Benjamins.

Moneglia, M. 2014. The variation of action verbs in multilingual spontaneous speech corpora: Semantic typology and corpus design. In T. Raso, and H. Mello (eds.), *Spoken corpora and linguistic studies*, 152-188. Amsterdam: John Benjamins.

Moneglia M., and Cresti E. 2006. C-ORAL-ROM Prosodic boundaries for spontaneous speech analysis. In Y. Kawaguchi, S. Zaima, and T. Takagaki (eds.), *Spoken language corpus and linguistics informatics*, 89-114. Amsterdam: John Benjamins.

Moneglia, M., Gagliardi, G., Panunzi, A., Frontini, F., Russo, I, and Monachini, M. 2012. IMAGACT: Deriving an action ontology from spoken corpora. In H. Bunt (ed.), *Proceedings of 8th ACL-ISO Workshop on interoperable semantic annotation (ISA-8)*, 42-47.

Moneglia, M., and Panunzi, A. 2007. Action predicates and the ontology of action across spoken language corpora: The basic issue of the SEMACT project. In M. Alcántara Plá, and T. Declerk (eds.), *Proceeding of the International Workshop on the Semantic Representation of Spoken Language (SRSL7)*, 51-58.

Moneglia, M., and Panunzi, A. 2011. Specification for the annotation of verb occurrences in the ImagAct project. Technical Report Draft, retrieved from http://lablita.dit.unifi.it/projects/IMAGACT/folder.2010-11-25.7365875310/ (last accessed: May 3, 2015).

Moneglia, M., Panunzi, A, Gagliardi, G., Monachini, M., Russo, I., De Felice, I., Khan, F., and Frontini, F. 2013. IMAGACT E-learning platform for basic action types. In *Proceedings of the 6th International Conference ICT for Language Learning (ICT4LL 2013)*, 85-90. Padova: Libreriauniversitaria.it.

Montefinese, M., Ambrosini, E., Fairfield, B., and Mammarella, N. 2013. Semantic memory: A feature-based analysis and new norms for Italian. *Behavior Research Methods* 45(2): 440-461.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., and Rizzolatti, G. 1997. Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology* 78: 2226-2230.

Murphy, M. L. 2010. *Lexical meaning*. Cambridge: Cambridge University Press.

Newell, K. M., Scully, D. M., Tenenbaum, F., and Hardiman, S. 1989. Body scale and the development of prehension. *Developmental Psychology* 22(1): 1-13.

Norman, D. A. 1988. *The Psychology of Everyday Things*. New York: Basic Books.

Norman, D. A. 1999. Affordance, conventions, and design. *Interactions* 6: 38-42.

Oztop, E., Arbib, M., and Bradley, N. 2006. The development of grasping and the mirror system. In A. Arbib (ed.), *Action to language via the mirror neuron system*, 397-423. Cambridge: Cambridge University Press.

Paivio, A. 1971. *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Panunzi, A., De Felice, I., Gregori, L., Jacoviello, S., Monachini, M., Moneglia, M., Quochi, V., and Russo, I. 2014. Translating action verbs using a dictionary of images: The IMAGACT ontology. In A. Abel, C. Vettori, and N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus*, 1163-1170. Bolzano: Eurac.

Pastra, K., Balta, E., Dimitrakis, P., and Karakatsiotis, G. 2011. Embodied language processing: A new generation of language technology. In *Proceedings of the AAAI 2011 International Workshop on "Language-action tools for cognitive artificial agents: Integrating vision, action and language"*, 23-29.

Pavese, A., and Buxbaum, L. J. 2002. Action matters: The role of action plans and object affordances in selection for action. *Visual cognition* 9(4/5): 559-590.

Pecher, D., and Zwaan, R. A. (eds.). 2005. *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge: Cambridge University Press.

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

Pietrandrea, P. 2002. Iconicity and arbitrareness in Italian Sign Language. *Sign Language Studies* 2(3): 296-321.

Puddu, M. 2000. *Ditzionàriu de sa limba e de sa cultura sarda*. Cagliari: Condaghes.

Pulvermüller, F. 2001. Brain reflections of words and their meaning. *Trends in Cognitive Sciences* 5: 517-524.

Pulvermüller, F. 2002. *The neuroscience of language*. Cambridge: Cambridge University Press.

Pulvermüller, F. 2005. Brain mechanisms linking language and action. *Nature Reviews Neurosciences* 6: 576-582.

Pulvermüller, F., Hauk, O., Nikulin, V. V., and Ilmoniemi, R. L. 2005. Functional links between motor and language systems. *European Journal of Neuroscience* 21: 793-797.

Pulvermüller, F., Shtyrov, Y., and Hauk, O. 2009. Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language* 110: 81-94.

Pustejovsky, J. 1991. The generative lexicon. *Computational Linguistics* 17(4): 409-441.

Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pustejovsky, J. 2013. Dynamic event structure and habitat theory. In R. Saurí, N. Calzolari, C. R. Huang, A. Lenci, M. Monachini, and J. Pustejovsky (eds.), *Proceedings of the 6th International Conference on Generative Approaches to the Lexicon*, 1-10.

Pylyshyn, Z. W. 1984. *Computation and cognition: Toward a foundation for cognitive science*. Cambridge: MIT press.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. 1985. *A comprehensive grammar of the English language*. London: Longman.

Radman, Z. (ed.). 2013. *The hand, an organ of the mind: What the manual tells the mental*. Cambridge, MA: MIT Press.

Raos, V., Umiltà, M., Murata, A., Fogassi, L., and Gallese, V. 2006. Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *Journal of Neurophysiological Studies* 95: 709-729.

Reed, E. S. 1988. Why do things look as they do? The implications of J. J. Gibson's "The ecological approach to visual perception". In G. Claxton (ed.), *Growth points in cognition*, 90-114. London: Routledge.

Ricciardi, E., Furey, M., Gobbini, M., Guazzelli, M., Haxby, J., and Pietrini, P. 2004. Tactile processing of different object categories involves extrastriate visual cortical areas in the human brain. In *Proceedings of EuroHaptics 2004*, 293-298.

Riddoch, M., Edwards, M., Humphreys, G., West, R., and Heafield, T. 1998. Visual affordances direct action: Neurophysiological evidence from manual interference. *Cognitive Neuropsychology* 15: 645-683.

Riordan, B., and Jones, M. N. 2011. Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science* 3(2): 1-43.

Rizzi, L. 1988. Il sintagma preposizionale. In L. Renzi, G. Salvi, and A. Cardinaletti (eds.), *Grande grammatica italiana di consultazione*, Vol. I, 507-531. Bologna: Il Mulino.

Rizzolatti G., and Arbib M. A. 1998. Language within our grasp. *Trends in Neurosciences* 21: 188-194.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. 1988. Functional organization of inferior area 6 in the macaque monkey. II: area F5 and the control of distal movements. *Experimental Brain Research* 71(3): 491-507.

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

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131-141.

Rizzolatti, G., and Luppino, G. 2001. The cortical motor system. *Neuron* 31: 889-901.

Rizzolatti, G., and Sinigaglia, C. 2006. *So quel che fai. Il cervello che agisce e i neuroni specchio*. Milano: Cortina.

Rome, E., Hertzberg, J., and Dorffner, G. (eds.) 2008. *Towards affordance-based robot control: International seminar, Dagstuhl Castle, Germany, June 5-9*. Berlin: Springer.

Ruimy, N., Monachini, M., Gola, E., Calzolari, N., Del Fiorentino, M. C., Ulivieri, M., and Rossi, S. 2003. A computational semantic lexicon of italian: SIMPLE. *Linguistica Computazionale* 18-19: 821-864.

Russo, I., De Felice, I., Frontini, F., Khan, F., and Monachini, M. 2013. (Fore)seeing actions in objects: Acquiring distinctive affordances from language. In B. Sharp, and M. Zock (eds.), *Proceedings of NLPCS 2013 - 10th International Workshop on Natural Language Processing and Cognitive Science*, 151-161.

Russo, I., Frontini, F., De Felice, I., Khan, F., and Monachini, M. 2013. Disambiguation of basic action types through nouns' Telic Qualia. In R. Saurí, N. Calzolari, C. R. Huang, A. Lenci, M. Monachini, and J. Pustejovsky (eds.), *Proceedings of the 6th International Conference on Generative Approaches to the Lexicon*, 70-75.

Şahin, E., Cakmak, M., Dogar, M.R., Ugur, E., and Ucoluk, G. 2007. To afford or not to afford: A new formalization of affordances towards affordance-based robot control. Adaptive Behavior 15(4): 447-472.

Şahin, E., and Erdogan, S. T. 2009. Towards linking affordances with mirror/canonical neurons. In *Proceedings of The 24th International Symposium on Computer and Information Sciences*, 397-404.

Salvi, G., and Vanelli, L. 2004. *Nuova grammatica italiana*. Bologna: Il Mulino.

Santos, A., Chaigneau, S. E., Simmons, W. K., and Barsalou, L. W. 2011. Property generation reflects word association and situated simulation. *Language and Cognition* 3(1): 83-119.

Sartori, L., Straulino, E., and Castiello, U. 2011. How objects are grasped: The interplay between affordances and end-goals. *PloS ONE* 6(9): e25203. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0025203 (last accessed: May 3, 2015).

Segalowitz, N. 2001. On the evolving connections between psychology and linguistics. *Annual Review of Applied Linguistics* 21: 3-22.

Shaw, R. E., Turvey, M. T., and Mace, W. 1982. Ecological psychology: The consequence of a commitment to realism. In W. Weimer, and D. Palermo (eds.), *Cognition and the symbolic Processes*, II: 159-226. Hillsdale: Lawrence Erlbaum Associates.

Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., and Just, M. A. 2011. Commonality of neural representations of words and pictures. *NeuroImage* 54: 2418-2425.

Simmons, W. K., Hamann, S. B., Harenski, C. N., Hu, X. P. and Barsalou, L. W. 2008. fMRI evidence for word association and situated simulation in conceptual processing. *Journal of Physiology-Paris* 102: 106-119.

Snow, R. E. 1992. Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist* 27: 5-32.

Spang-Hanssen, E. 1963. *Les prépositions incolores du français moderne*. Copenhague: Gads.

Steedman, M. 2009. Foundations of Universal Grammar in planned action. In M. Christiansen, C. Collins, and S. Edelman (eds.), *Language Universals*, 174-199. Oxford: Oxford University Press.

Sun, J., Moore, J. L., Bobick, A., and Rehg, J. M. 2010. Learning visual object categories for robot affordance prediction. *The International Journal of Robotics Research* 29(2-3): 174-197.

Symes, E., Ellis, R., and Tucker, M. 2007. Visual object affordances: Object orientation. *Acta Psychologica* 124: 238-255.

Taub, S. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge, MA: Cambridge University Press.

Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., and Moro, A. 2008. Negation in the brain: Modulating action representations. *NeuroImage* 43: 358-367.

Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T., and Baldassarre, G. 2013. Theories and computational models of affordance and mirror systems: An integrative review. *Neuroscience and Biobehavioral Reviews* 37: 491-521.

Tremblay, P., Sato, M., and Small, S. L. 2012. TMS-induced modulation of action sentence priming in the ventral premotor cortex. *Neurpsychologia* 50: 319-326.

Tucker, M., and Ellis, R. 1998. On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance* 24: 830-846.

Tucker, M., and Ellis, R. 2001. The potentiation of grasp types during visual object categorization. *Visual Cognition* 8: 769-800.

Tucker, M., and Ellis, R. 2004. Action priming by briefly presented objects. *Acta Psychologica* 116: 185-203.

Turvey, M. T. 1992. Affordances and prospective control: An outline of the ontology. *Ecological Psychology* 4(3): 173-187.

Turvey, M. T., Shaw, R. E., Reed, E. S., and Mace, W. M. 1981. Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn. *Cognition* 9: 237-304.

Ugur, E., Oztop, E., and Şahin, E. 2009. Affordance learning from range data for multi-step planning. In *Proceedings of the 9th international conference on epigenetic robotics*, 177-184. Lund University Cognitive Studies.

Varvara, R., and Ježek, E. 2014. Semantic role annotation of instrument subjects. In R. Basili, A. Lenci, and B. Magnini (eds.), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and the Fourth International Workshop EVALITA 2014*, 384-388. Pisa: Pisa University Press.

Vermaas, P., Kroes, P., van de Poel, I., Franssen, M., and Houkes, W. 2011. A philosophy of technology: From technical artefacts to sociotechnical systems. *Synthesis Lectures on Engineers, Technology and Society* 6(1): 1-134.

Vigliocco, G., and Vinson, D. P. 2007. Semantic representation. In G. M. Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 195-215. Oxford: Oxford University Press.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., and Cappa, S. F. 2011. Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews* 35: 407-426.

Vinson, D. P., and Vigliocco, G. 2008. Semantic feature production norms for a large set of objects and events. *Behavior Research Methods* 40(1): 183-190.

Warren, W. 1984. Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 5(10): 683-703.

Warren, W., and Whang, S. 1987. Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance* 13(3): 371-383.

Weinrich, H. 1988. Preposizioni incolori? Sulle preposizioni, franc. *de* e *a*, ital. *da*. In H. Weinrich, *Lingua e linguaggio nei testi*, 139-180. Milano: Feltrinelli.

Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., and Puce, A. 2004. Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *NeuroImage* 22: 277-288.

Wierzbicka, A. 1988. Oats and wheat: The fallacy of arbitrariness. In A. Wierzbicka, *The Semantics of Grammar*, 499-560. Amsterdam: John Benjamins.

Wilcox, S. 2004. Conceptual spaces and embodied actions: Cognitive iconicity and signed languages. *Cognitive Linguistics* 15(2): 119-147.

Witt, J. K., Proffitt, D. R., and Epstein, W. 2004. Perceiving distance: A role of efforts and intent. *Perception* 33: 577-590.

Witt, J. K., Proffitt, D. R., and Epstein, W. 2005. Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance* 31: 880-888.

Wu, L., and Barsalou, L. W. 2009. Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica* 132(2): 173-189.

Ye, L. 2010. Placing objects in the context of goal-directed actions: Cultural differences between Chinese and American students in the perception of multiple affordances for objects. Master Thesis, Miamy University, Oxford, Ohio. Retrieved from https://etd.ohiolink.edu/ (last accessed: May 3, 2015).

Yeh, W., and Barsalou, L. W. 2006. The situated nature of concepts. *American Journal of Psychology* 119: 349-384.

Zwaan, R. A., and Taylor, L. J. 2006. Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General* 135: 1-11.