

# Semantic properties of word associations to Italian verbs

Annamaria Guida & Alessandro Lenci

This work is concerned with an investigation of verb associations, i.e., the words spontaneously called to mind in response to a given stimulus verb. We performed an elicitation task where native speakers were asked to spontaneously list semantic associations for a list of Italian verbs. Starting from the assumption that the associations reflect highly salient linguistic and conceptual features of the verbs, the investigation is directed toward specifying the structural and conceptual types of associations by distinguishing and quantifying the relationships between stimuli and responses\*.

## *1. Introduction*

This paper presents an analysis of a collection of semantic associations evoked by a list of Italian verbs in a word association experiment.

We define WORD ASSOCIATIONS as those words spontaneously called to mind in response to a given stimulus word. Word associations have been of interest to psycholinguistics for decades. They have been used over the years as a tool to investigate the mechanisms underlying semantic memory, giving the researchers a relatively transparent measure of the semantic information that is normally accessed when a word is heard or read. For this reason, they have facilitated the development of empirically grounded models of lexical-semantic knowledge. Specifically, they have been used to address research questions that range from word recognition in semantic priming experiments (cf. McNamara 2005) and memory research (Nelson et al. 1997; Nelson & Zhang 2000) to the development of semantic networks (Plaut 1995; Burgess 1998).

Collections of word associations – referred to as ASSOCIATION NORMS – have been established for many languages. Association norms are created by defining a set of target stimuli (depending on the purpose of the norms, controlling for e.g. the number of syllables, frequency, etc.) and asking participants to provide the first word(s) that come to mind when presented with the stimuli. Then, the results are collapsed across participants, quantifying over the number of response tokens for each stimulus-response pair. Along with the type

of response, frequency of response is considered to be an essential index, a measure of the strength of semantic relation between words. Following an idea originally suggested by the pioneering British psychologist Francis Galton in 1880, the first association norms were collected by Kent & Rosanoff (1910) on the base of a list of 100 stimulus words including common nouns and adjectives, and 1000 participants being involved. The Kent & Rosanoff stimuli were then translated into several languages (e.g. German: Russell & Meseck 1959; Russell 1970), allowing for the collection of parallel association norms for languages other than English. In spite of the numerous advantages of using as stimuli the translations of Kent & Rosanoff stimulus words (first of all, the possibility of cross-language comparisons), a by-product of focussing on Kent & Rosanoff words was that much of the research was restricted to (highly frequent) nouns and adjectives. To overcome this limit, another collection was assembled by Palermo and Jenkins (1964), comprising associations for 200 words across various parts-of-speech. The first attempt to collect association norms on a larger scale was the Edinburgh Association Thesaurus (Kiss et al. 1973). In a similar vein, the association norms from the University of South Florida (Nelson et al. 1998) were collected over the course of more than 20 years. Their goal was to obtain the “largest database of free associations ever collected in the United States available to interested researchers and scholars”. More than 6,000 participants produced nearly three-quarters of a million responses to 5,019 stimulus words. Smaller sets of association norms have also been collected for example for Dutch (Lautenslager et al. 1986), French (Ferrand & Alario, 1998), Italian (Peressotti et al. 2002), Spanish (Fernandez et al. 2004), and German (Schulte im Walde & Melinger 2005; Melinger & Weber 2006).

After this overview of previous work related to word associations collection, we would like to provide a brief description of some studies with direct relevance to the present paper. Indeed, in spite of the great amount of word associations data that are available in several languages, few investigations have studied the properties of the associations in depth. In early work on association norms, Clark (1971) identified potential relations between stimulus words and their associations on a theoretical basis, categorising stimulus-response relations into sub-categories such as synonymy, antonymy, selectional preferences etc. Heringer (1986) asked his participants to provide question words (e.g., *wer* ‘who’, *was* ‘what’) as associations to the 20 German verbs selected as stimuli, in order to investigate the valency behaviour of the verbs. But the more extensive investiga-

tions of the properties of semantic associates can be found in recent works by Schulte im Walde and colleagues. Schulte im Walde & Melinger (2005) collected norms for 330 German verbs and conducted several empirical analyses on them, providing detailed insights into the semantic relations and the functional properties instantiated by the elicited associates. On the other hand, Melinger & Weber (2006) collected a similar database of associations for a list of 409 German nouns. Armed with these two datasets of association norms for verbs and nouns, a series of following studies presented various extensions of the basic analysis and several application scenarios. Interestingly, their work is situated between the psychological and the computational lines of research: not only do they make use of large-scale computational resources and methods to analyse the association norms, but their insights are also thought to contribute to both cognitive and computational linguistic modelling. For instance, Schulte im Walde (2006a; 2008) relied on the collected verb association norms to investigate whether word associations can help us to identify salient features for semantic verb classes. She applied a cluster analysis to the verbs, as based on the associations, and validated the resulting verb classes against standard approaches to semantic verb classes. Melinger et al. (2006) took the noun associations as input to a soft clustering approach, in order to determine the various noun senses of ambiguous stimulus nouns. Finally, Schulte im Walde & Melinger (to appear) performed a detailed analysis of corpus co-occurrences distribution of semantic associations. They relied on the assumption that there is a high correlation between associative strength in association norms and word co-occurrence in language corpora (see Spence & Owens 1990), and claimed that this fact can profitably be exploited in Natural Language Processing (NLP) to build semantic representations based on distributional data, and to determine the types of semantic relationships relevant for computational lexicons. Bringing together all their analyses, word association data have been proved to be useful both in psycholinguistics and in NLP, and have been profitably used to address research questions concerning semantic relatedness from several perspectives.

## *2. Goal of the paper*

This paper uses word associations as the basis for an investigation of semantic properties of Italian verbs. Our central assumption is that associations to verbs model salient aspects of verb meaning and

that they should therefore represent a good basis for an investigation of verb semantic representation (cf. among others: Tanenhaus et al. 1989; McKoon & Ratcliff 1992; Plaut 1995; McRae & Boisvert 1998). To our knowledge, no data on associations evoked by Italian verbs are available so far (indeed, the aforementioned association norms collected by Peressotti et al. 2002 do not provide norms for verbs: their stimuli consist of 289 nouns, 4 adjectives and 3 adverbs). Therefore, our data collection is intended to provide a starting point to fill a gap in available behavioural data about the meaning of Italian verbs.

The primary aim of this work is to conduct an in depth examination of the properties of stimulus-response pairs, in order to probe the range of semantic information that is accessed when a verb is presented in isolation, i.e., when a linguistic or extra-linguistic context is absent. More specifically, this work addresses the need for an analysis of the different kinds of information derived from the association test: which kinds of semantic information are automatically accessed when a verb is heard/read? For example, which kinds of semantic verb relations are instantiated by verb responses, and which kinds of linguistic functions are instantiated by noun responses? Indeed, for a stimulus verb to elicitate a particular response, it is necessary that the semantic information underpinning that stimulus-response relationship is accessed when the stimulus verb is presented. For example, if *ignore* evokes *know*, this suggests that when *ignore* is heard/read, information about its antonym is accessed. If we find that *drink* evokes *water* or *wine*, it suggests that the information accessed include nouns that could function as its prototypical patients. Distinguishing between different types of lexical semantic relations can help us deepen our insight into the various mechanisms that seem to operate when a speaker is presented with such a task, as well as to provide further evidence about the organization of semantic lexical knowledge.

We would like to close this section with an important remark on a methodological issue regarding word association tasks. As aforementioned, these tasks require participants to provide associations to words that are presented out of context. According to a longstanding tradition, tasks aimed at investigating lexical-semantic representations (e.g. feature generation tasks; see McRae et al. 2005) make use of decontextualized words as stimuli. Indeed, it is reasonable to assume that certain aspects of information within word meaning representation are more or less salient, and that the most salient aspects will be automatically activated in the absence of a supporting context. However, this assumption is worth being reviewed and debated in the light of the recent discussion in cognitive science on the situ-

ated nature of conceptualization (Glenberg & Kaschak 2002; Barsalou 2005; Wu & Barsalou *submitted*). According to the situated cognition view, conceptual representations, rather than being abstract, decontextualized and stable, are grounded to some extent on perception and action. Wu & Barsalou collected evidence from feature generation data showing a strong correlation between properties generated by participants explicitly instructed to use mental images and the properties produced by participants that did not receive such an instruction. These results are interpreted as supporting the view that participants generate properties of a concept by “running” perceptual simulations of its instances. Moreover, Wu & Barsalou show that an average of 25% of the properties produced by their participants are related to aspects of the prototypical contextual setting of the concept instances, such as typical actions and locations, entities co-occurring in the same context, etc. A situated view of concept representation is even more reasonable for the representation of the meaning of verbs. Whereas nouns denoting objects can be understood in isolation, events are relational in nature: it has been argued that the lexical-semantic representations for events contain a sort of “core” meaning (the event denoted) and, strictly related to this meaning component, the representation of the (number and kind of) participants involved in the event (e.g., the verb *beat* implies hitting repeatedly, and thematic roles that refer to the roles played by the verb in terms of “who did what to whom”), as well as various aspects of the event “scenario” (e.g. its location, related events, etc.). If this is true, performing a word association test on verbs, we expect to find in our data a confirmation of Wu and Barsalou’s results, i.e., we expect to find among our associates a considerable number of words referring to aspects of the prototypical contextual setting in which the event denoted by the verb may happen.

The remainder of this paper is organised as follows. In Section 3, the collection of associations, performed in a web experiment, is discussed in some detail. A series of empirical linguistic analyses of the data are described in Section 4. We conclude with a discussion of the results of this work as well as some open issues in Section 5.

### *3. Word association experiment*

This section introduces the word association experiment we performed in order to collect associations provided by speakers. Details on the material selected as stimuli, elicitation method and data pre-processing are described respectively in Sections 3.1, 3.2. and 3.3.

### 3.1 Material

A set of 312 Italian verbs was selected and used for the experiment. The verbs for which associations were collected were chosen to cover a broad range of verb types with respect to various parameters.

First, the verbs were drawn from a variety of semantic classes, since “semantic verb classes generalize over verbs according to their semantic properties, i.e., they capture large amounts of verb meaning without defining the idiosyncratic details for each verb.” (Schulte im Walde 2006b: 159). The verbs were manually classified into 18 concise semantic verb classes. Appendix A presents a full list of the verbs included in the study and their classification into semantic classes. Examples for semantic verb classes are MOTION verbs such as *andare* ‘go’, TRANSFER OF POSSESSION verbs such as *dare* ‘give’, COMMUNICATION verbs such as *dire* ‘say/tell’. Some of the classes are divided into sub-classes according to salient semantic and syntactic distinctions. For instance, following Talmy (1985) and subsequent research (Jackendoff 1990; Levin and Rappaport Hovav 1992; Slobin 1996), verbs that express motion can be decomposed into basic conceptual elements. MOTION verb meanings normally contain a “path” (the direction of the movement) and the manner of movement (e.g., *walking* vs. *running*). Some motion verbs encode the DIRECTION OF MOTION (e.g., *entrare* ‘go into’, *uscire* ‘go out’, *cadere* ‘fall’) and need optional adverbial phrases to express manner of motion (*entrare correndo* ‘enter running’). Other verbs encode MANNER OF MOTION (e.g., *camminare* ‘walk’, *correre* ‘run’, *saltare* ‘jump’, *rotolare* ‘roll’), using prepositional phrases or adverbs to indicate the direction (*correre in casa* ‘run into the house’). These verbs expressing manner of motion can be sub-divided into finer labels, e.g., including a sub-class for verbs expressing MANNER OF MOTION USING A VEHICLE. The inclusion of any verb in any particular verb class was achieved with reference to prior verb classification work, and in particular this classification closely follows the proposal made by Levin (1993) for English & Schulte im Walde (2006b) for German. Nonetheless, distinctions between classes were sometimes hard to make, and this is reinforced by the fact that classes may have several verbs in common, according to traditional classification by Levin (1993). For instance, many verbs cannot be unequivocally classified as either COGNITION or COMMUNICATION verbs: indeed the verb *giudicare* ‘judge’ refers both to the mental activity of judging and to the action of articulating one’s judging.

Second, the stimuli include both high-frequency and low-frequency verbs. Frequencies were computed from *La Repubblica* Corpus, a 380 million word newspaper corpus (Baroni et al. 2004). The verbs showed corpus frequencies between 28 (*faxare* 'fax') and 963,273 (*fare* 'do'). They were drawn from 11 different frequency ranges. Thus, the verbs were chosen to cover a wide range of familiarity, although, of course, participants should be reasonably familiar with all the verbs to be able to produce useful information about them. The degree of abstractness of semantic content is also strictly correlated with verb frequency: for example, we can contrast a very generic MOTION verb such as *andare* 'go' (highly frequent) and a very specific concept like *marciare* 'march', that exhibits a far lower corpus frequency.

Moreover, the verbs for this experiment were chosen to show a highly variable degree of polysemy. They vary from unique-sense verbs like *arrossire* 'blush', to verbs with a variety of polysemic senses like *ordinare*, for which many different senses can be listed: roughly, 'put in order, arrange', 'regulate', 'tidy up', 'marshal', 'decree, dispose', 'command', 'prescribe'. As a consequence, these verbs are ambiguous with respect to class membership. An example of class membership ambiguity is the verb *sostenere*: it swings from the sense 'hold up, sustain', and therefore may be seen as a POSITION verb, to the sense 'support, help', which would lead the verb toward the SUPPORT verbs class, up to the sense 'believe, think', which would place this verb among the COGNITION verbs. In these cases, we arbitrarily assigned the verb to a particular class according to one of its senses (so that, e.g., *sostenere* belongs to the COGNITION class). However, this is not particularly crucial for our analysis. Including polysemous verbs in the list was only intended to provide us with the possibility to evaluate the effect of polysemy on word associations distribution. Relying on the assumption that associates represent a useful basis for understanding which kind of information is more salient to define the core meaning of the verbs, it can be interesting to discover which of the different senses of polysemous verbs will emerge as dominant from its associates, since this finding could provide us with some evidence of the salience of a particular word sense for native speakers.

The selected verbs describe different events or situation types, i.e. they express different actionality (Aktionsart) values (for an exhaustive discussion and taxonomy of event types in Italian verbal system cf. Bertinetto 1986, Bertinetto & Squartini 1995). For the purposes of the present discussion, it will be enough to take the category Aktionsart in the sense of the traditional four Vendlerian classes (sta-

tives, activities, achievements, accomplishments; cf. Vendler 1967), whose reciprocal delimitations are based on the features [ $\pm$ durative], [ $\pm$ dynamic], [ $\pm$ homogeneous]. Actionality has to do with the nature of the event type associated with a verbal predicate. Therefore, most verbs may have more than one actional classification. However, among the experiment stimuli some verbs are inherently stative (*sapere* ‘know’, *volere* ‘want’), others are instead inherently dynamic (*correre* ‘run’). Moreover, some verbs in this latter group are inherently telic (*rompere* ‘break’).

Finally, the verbs show a broad variation in the numbers of selected arguments. The minimum number of semantic argument is zero, in WEATHER verbs such as *piovere* ‘rain’; the largest number of semantic arguments is four, exhibited by verbs describing commercial transactions (within the semantic class of TRANSFER OF POSSESSION verbs) such as *vendere* ‘sell’.

### 3.2 Method

*Procedure:* The 312 verbs were divided in 6 separate experimental lists of 52 verbs each, so that every participant had to provide associations to 52 verbs. These lists were balanced for class affiliation and frequency ranges, that is each list contained verbs from each grossly defined semantic class and had equivalent overall verb frequencies distributions. The procedure used by Schulte im Walde & Melinger (2005) in their word association experiment on German verbs was closely followed. The experiment was administered over the Internet,<sup>1</sup> and we used the program developed by Schulte im Walde & Melinger (2005), after translating the instructions into Italian and adapting them, where necessary, to the new task. The program was compatible with most browsers and platforms. When participants loaded the experimental page, they were first asked for their biographical information, such as linguistic expertise, age, profession and region. The following page was loaded by the participants by clicking the button “Next” with the mouse. At this point, the participant was presented with the written instructions for the experiment and an example item with potential responses. They were instructed to not wait too long reflecting, but to type the first words that spontaneously come to their mind. They were also asked to type at most one word per line. Stimuli presentation began when participants clicked the button “Start”. In the actual experiment, each trial consisted of a verb (in the infinitive) presented in a box at the top of the screen. Below the verb was a series of blank lines where partici-

pants could list their associations. Participants had 30 sec. per verb to type as many associations as they could. After 30 sec., the verb disappeared from the screen and the program automatically advanced to the next trial (with a short pause of 2 sec. between a verb and the next one). When participants loaded the experimental page, one of the 6 verbs lists was randomly selected and, for each list, the order of the verbs was randomized through each participant, in order to avoid order effects. The experiment took approximately 30 min. At the end of the task the data were automatically saved to an individually named file.

*Participants:* Participants were recruited in different ways. For the most part they were recruited placing several links to the experiment on web sites, Internet forums etc. Moreover, our experiment was announced as part of the collection of web psycholinguistic experiments on the web laboratory “Portal for Psychological Experiments on Language” (<http://www.language-experiments.org/>). Experiment data sets were collected from January 31st to May 1st 2007. A total of 292 native Italian speakers participated in the experiment providing responses for at least 75% of verbs (we disregarded the other data sets, considering them as highly incomplete). The participants were between 46 and 54 for each experimental list.

### *3.3 Data preparation*

Each completed data set contains the background information of the participant, followed by the list of target verbs. Each target verb is paired with a list of associations in the order in which the participant provided them. In total, we collected 78,700 associations from 14,835 trials. Each trial elicited an average of 5,3 associate responses with a range of 0-15.

For the analyses to follow, we preprocessed all data sets in the following way: for each data set, we extracted only the first response each participant provided for each verb. That is to say, we extracted from each data set 52 couples stimulus-first response, disregarding all the associations provided by each participant after the first one. In this way, we considered only the first response for each trial, and therefore we collected in total 14,835 associations. For illustrative purposes, Table 1 lists all the associations provided as first response by each participant for the stimulus verb *chiudere* ‘close’. The associations are provided with their English translations and with their frequencies (the number of participants who produced the response).

**Table 1.** List of associations provided for the verb *chiudere*.

<i>chiudere</i> ‘close’		
<i>porta</i>	‘door’	37
<i>serrare</i>	‘shut’	8
<i>aprire</i>	‘open’	2
<i>bocca</i>	‘mouth’	1
<i>chiuso</i>	‘closed’	1
<i>sbattere</i>	‘slam’	1
<i>smettere</i>	‘quit, cease’	1
<i>storia</i>	‘affair’	1

The choice of selecting only the “first word that comes to mind” has a twofold motivation. First, investigations into the reliability of associations have shown that the first response is at least sufficient and possibly superior to subsequent responses (McEvoy & Nelson 1982; Nelson et al. 2000). This assumption is due to the finding that if we consider the fastest responses, that is, responses given without reflexion, there is a much higher degree of consistency in the results, whereas later responses are often idiosyncratic, precious responses as well as personal recollections. Second, the so-called association chaining, i.e., that the *n*th response is associated to the (*n*-1)th response rather than to the stimulus, is a phenomenon that somehow contaminates later responses (McEvoy & Nelson 1982; Nelson et al. 2000) and makes their analysis much more complex.

For these reasons, we decided to narrow the field of our analysis in this paper to the first associates. Nevertheless, we are interested in whether and to what extent the analysis of all the responses could provide a richer picture of the semantics of the target by pointing to additional meaning component (as recently suggested by Schulte im Walde and Melinger, *to appear*). Therefore, it is our plan to submit the full set of associates we collected (78,700 responses) to further analyses.

#### 4. Analysis of experiment data

We are interested in the types of relationship between the associates and the verb stimuli. To address this specific point, we conducted the following three analyses:

1. in a preparatory step, we classified the responses with respect to part-of-speech tags (Section 4.1);
2. for each verb associate, we then tried to determine the semantic relation between the target and response verbs (Section 4.2);
3. finally, for each noun associate, we investigated the kinds of

semantic functions that are realized by the noun with respect to the target verb (Section 4.3).

#### *4.1. Morpho-syntactic analysis: Part-of-Speech tagging*

As a first step for the linguistic analysis of our responses, the associates have been distinguished with respect to the part-of-speech (henceforth: PoS) tags. Each associate of the target verb was assigned to its (possibly ambiguous) PoS. A number of words turned out to be ambiguous between various PoS tags. Ambiguity arose especially in the case of nominalised adjectives, where the experiment participant could have been referring either to the adjective (*caldo* ‘warm’) or the noun (*caldo* ‘warmth’), in the case of homograph pairs noun/verb (such as *potere* ‘power/be able’, *dovere* ‘duty/must, have to’) and in the case of adjectives/adverbs pairs (such as *veloce* ‘quick’). When a response word was grammatically ambiguous, information provided by the stimulus word was used to decide the correct interpretation. Some appeal to intuition was made in deciding whether, given the stimulus, one or the other of the possible usages of the response word was likely to be the dominant one. For instance, *potere* was provided as response to *bramare* ‘crave’, and it is clear in this case that *potere* was taken as a noun (meaning ‘power’), indicating the patient of the stimulus verb.

Having assigned PoS tags to the responses, we could distinguish and quantify the morpho-syntactic categories of the associates. Among the 14,835 responses, nouns, verbs, adjectives, adverbs, pronouns, conjunctions, prepositions, interjections were found. Anyway, nouns, verbs, adjectives and adverbs cover more than 99% of all the responses. Therefore, the actual classification is obtained by grouping together all the other categories in a unique category.

As a result of this first analysis, we can specify the frequency distributions of the PoS tags for each verb and in total. Participants provided noun associates in the majority of token instances,<sup>2</sup> 54.4%; verbs were given in 38.8% of the responses, adjectives in 3.4%, adverbs in 2.5%; residual categories cover 0.9% of the responses.

The PoS distribution for responses is correlated with target verb frequency. The rate of verb and adverb responses is positively correlated with target verb frequency, Pearson’s  $r=0.24$  for verbs and  $r = 0.14$  for adverbs, while the rate of noun is inversely correlated with verb frequency, Pearson’s  $r = -0.28$ . These results are highly consistent with correlations observed by Schulte im Walde & Melinger (2005) in their word association experiment on German verbs.

The distribution of responses over PoS also varies across verb classes, as we can see from Table 2.

**Table 2.** Distribution of responses over part-of-speech across verb classes.

Verb class	N%	V%	Adj%	Adv%	Others%
Cognition	48	45.1	4.2	1.8	0.9
Desire	48.5	47.3	2.6	0.7	0.9
Transfer of possession – Giving	51.3	44.8	1.7	1.2	1
Transfer of possession – Obtaining	54.8	39.2	1.2	4.6	0.2
Motion	58.4	33.8	2	5.6	0.2
Emotion	60.2	28.8	6.6	1.7	2.7
Perception	52.7	43.5	1.3	2	0.5
Communication	51	42.9	1.5	3.8	0.8
Teaching and learning	50.6	46.3	1.2	1.6	0.3
Position - Be in Position	43.8	42	9.6	4.6	0
Position - Bring into Position	37.4	60.3	0.8	1.4	0.1
Creation	54	42.8	1.6	0.5	1.1
Change	48.8	41.2	7.8	2.1	0.1
Consumption	55.2	38.1	4.2	1.7	0.8
Elimination	45.2	51.9	1.2	0.6	1.1
Weather	80.2	12.5	5.3	1.8	0.2
Measure	68.6	23.4	4.3	2.1	1.6
Verbs involving the body – Dressing	56.9	33.5	7.9	1.7	0
Verbs involving the body – Bodily Processes	71.4	21.8	5.1	0.8	0.9
Verbs involving the body – Gestures	56.4	26.6	5.4	11.2	0.4
Verbs involving the body – Damage to the Body	66.2	31.3	2.1	0.2	0.2
Performative	55.9	39.2	1.6	1.2	2.1
Support	62	33.3	1.4	1.4	1.9
Aspect	28.3	62.9	0.6	5.7	2.5
<b>TOTAL</b>	<b>54.4</b>	<b>38.8</b>	<b>3.4</b>	<b>2.5</b>	<b>0.9</b>

For example, ASPECT verbs received much more verb responses (62.9%) than noun responses (28.3%), and also BRING INTO POSITION verbs (verb responses: 60.3%; noun responses: 37.4%) show a similar distribution, whereas WEATHER verbs received 80.2% of noun responses and only 12.5% of verb responses and, following the same tendency, BODILY PROCESSES verbs received 71.4% of noun associates and 21.8% of verb associates. Tables 3 and 4 show the associations provided respectively for the ASPECT verb *terminare* ‘end’ (verbs: 78.7%, nouns: 21.2%) and for the BODILY PROCESSES verb *pungere* ‘prick, sting’ (verbs: 3.7%, nouns: 96.2%):

**Table 3.** List of associations provided for the verb *terminare*.

<i>terminare</i> 'end'		
<i>finire</i>	'finish'	32
<i>fine</i>	'end'	4
<i>concludere</i>	'conclude'	3
<i>chiudere</i>	'close'	1
<i>morire</i>	'die'	1
<i>calcolo</i>	'calculation'	1
<i>capolinea</i>	'end of the line'	1
<i>esperimento</i>	'experiment'	1
<i>partita</i>	'match'	1
<i>sollievo</i>	'relief'	1
<i>supplizio</i>	'torment'	1

**Table 4.** List of associations provided for the verb *pungere*.

<i>pungere</i> 'prick, sting'		
<i>ape</i>	'bee'	16
<i>ago</i>	'needle'	15
<i>dolore</i>	'pain, ache'	7
<i>zanzara</i>	'mosquito'	6
<i>insetto</i>	'insect'	2
<i>dito</i>	'finger'	2
<i>pizzicotto</i>	'pinch'	1
<i>puntura</i>	'prick, sting'	1
<i>rosa</i>	'rose'	1
<i>spillo</i>	'pin'	1
<i>stuzzicare</i>	'prod, poke'	1
<i>punzecchiare</i>	'tease'	1

The distribution varies largely also across verbs within the same class. For instance, within the MOTION class, participants provided 93.6% of noun associates for *pedalare* 'cycle', whereas noun associates constitute only 20.4% of the total for the verb *andare* 'go' and 17.7% for the verb *fuggire* 'run away, flee'. This fact is consistent with the correlation between verb frequency and number of noun associates which we have pointed out above. Given the well-known inverse correlation between frequency and semantic specificity, it seems that specific verbs tend to trigger noun associates, more than generic verbs. We might even conjecture that a verb denoting a highly specific event or situation (e.g. cycling: see Table 5) tends to trigger nouns referring to entities typically participating in the event (e.g. bike).

**Table 5.** List of associations provided for the verb *pedalare*.

<i>pedalare</i> 'cycle'		
<i>bicicletta</i>	'bike'	38
<i>fatica</i>	'fatigue'	3
<i>correre</i>	'run'	2
<i>muoversi</i>	'move'	1
<i>noia</i>	'boredom'	1
<i>pedali</i>	'pedals'	1
<i>sudore</i>	'sweat'	1

Concerning adjectives, the verb class that presents the highest percentage of adjective responses is BE IN POSITION class, with 9.6%, followed by VERBS INVOLVING THE BODY-DRESSING (7.9%) and by CHANGE verbs (7.8%). The verb that turns out to be the richest in adjective responses is *impallidire* 'turn pale'. In this case, adjectives constitute 40.4% of the whole set of associations (Table 6).

**Table 6.** List of associations provided for the verb *impallidire*.

<i>impallidire</i> 'turn pale'		
<i>bianco</i>	'white'	16
<i>sbiancare</i>	'turn white'	9
<i>paura</i>	'fright'	9
<i>viso</i>	'face'	3
<i>pallido</i>	'pale'	2
<i>guance</i>	'cheeks'	2
<i>arrossire</i>	'blush'	1
<i>blocco</i>	'block'	1
<i>innervosire</i>	'irritate'	1
<i>sconvolto</i>	'deranged'	1
<i>svenire</i>	'faint'	1
<i>biancore</i>	'whiteness'	1

Many verbs with a high percentage of adjectives among associates are, exactly as *impallidire*, deadjectival verbs of change,<sup>3</sup> formed with a prefix from their root adjectives: *ingrassare* 'fatten', *allungare* 'lengthen', *allargare* 'enlarge', etc.<sup>4</sup>

Finally, the verb class that turns to be the richest in adverb responses is BODILY GESTURES class. In this case, however, the high number of adverbial response tokens (27) corresponds to only 2 adverbial response types (namely, the adverbs *si* 'yes' and *ok* elicited by the verb *annuire* 'nod'). On the contrary, a class in which adverbial responses are quite numerous and distributed over the whole class is MOTION verbs class. Motion verbs evoked many different adverb types, mostly direction adverbs (for instance: *tornare* 'come back' evoked *indietro* 'back'; *volare* 'fly' evoked *via* 'away') and manner adverbs

(*correre* ‘run’ – *veloce* ‘quick, fast’; *camminare* ‘walk’ – *piano* ‘slowly’). As we could expect, motion verbs encoding the direction of motion evoked a great majority of direction adverbs and, on the other hand, MANNER OF MOTION verbs evoked mostly manner adverbs.

#### *4.2. Semantic relations of verb associates*

In a second analysis we investigated, for each verb associate, the types of semantic relation between the target verb and response verb using as a basis for classification the lexical semantic taxonomy WordNet (Miller et al. 1990; Fellbaum 1998) and its Italian counterpart ItalWordNet (IWN; see Roventini et al. 2000). WordNet is inspired by psycholinguistic research on lexical memory and it is structured around the notion of synset, i.e., set of synonymous word meanings, with basic semantic relations encoded between the synsets. The relations encoded in WordNet can all be thought of pointers or labeled arcs from one synset to another. Word meanings, linked by semantic relations, form a complex semantic network; therefore, in WordNet a word is basically described by means of its relations with other word meanings. The underlying assumption is that knowing where a word is located in that network is an important part of knowing the word’s meaning. In WordNet lexical entries are separated according to their syntactic category membership. The relations encoded in WordNet between verb synsets are: synonymy, antonymy, troponymy,<sup>5</sup> entailment, cause.<sup>6</sup> Based on these relations, we could distinguish between the different kinds of verb associations elicited from speakers. Since data from word association experiments are considered as a good source of evidence of the organisation of speakers’ mental lexicon, we will discuss some results coming from our analysis of semantic verb relations and we will test whether the psycholinguistic assumptions underlying the WordNet model fit well with the stimulus-response pairs.

Our analysis proceeds as follows. For each pair of target and response verbs, we look up whether any kind of semantic relation is defined between any of the synsets the verbs belong to in ItalWordNet. In all the cases in which either member of the verb pair is not present in ItalWordNet, or in the cases in which they are both in ItalWordNet but there is no relation between their synsets, we manually labeled the semantic relation between the target and the response. We then calculated the frequency of each semantic relation: for instance, since 15 participants provided the association *terminare* ‘finish’ for its synonym *finire* ‘end’, the synonymy relation is assigned 15 as its frequency value.

As a result of this classification, we can specify the frequency distributions of the semantic relations for each verb individually, and also as a sum over all verbs. If we consider only verb-verb pairs we obtain the overall picture described in Table 7. The labels indicate the relation between the response and the stimulus.

**Table 7.** The distribution of semantic relations over the set of verb responses.

Verb semantic relations	Examples stimulus/response	Percentage on verb responses
Troponymy (superordinate)	<i>scrutare/vedere</i> ‘peer/see’	22.8
Troponymy (subordinate)	<i>parlare/chiacchierare</i> ‘talk/chat’	5.9
Troponymy (coordinate)	<i>correre/camminare</i> ‘run/walk’	11.7
Synonymy	<i>modificare/cambiare</i> ‘modify/change’	38.3
Antonymy	<i>cominciare/finire</i> ‘begin/end’	4.5
Cause	<i>ansimare/correre</i> ‘pant/run’	8.2
Entailment	<i>russare/dormire</i> ‘snore/sleep’	2.6
Unknown cases	<i>dire/fare</i> ‘say/do, make’	6

Among verb associates, troponymy and synonymy have turned out to be the most frequent relations linking stimuli and responses, and this finding has been assumed to point to and confirm the psychological salience of these relations. Moreover, an overview of the results reveals that the distribution of semantic relations also varies by verb class. For example, according to our results the most frequent association for *cominciare* ‘begin’, is its synonym *iniziare* while for *vendere* ‘sell’, it is its converse *comprare* ‘buy’, and for *scrutare* ‘peer’, its superordinate *guardare* ‘look at’. Although synonymy can be considered as a highly pervasive relation over all the verb classes, there are particular verbs for which synonymic responses cover the vast majority of associates: *terminare* ‘stop’ and *cominciare* ‘begin’, for instance, evoked respectively 85% and 78% of synonymic responses.

Troponymy relation seems to fit a large number of association pairs. Specifically, on the basis of troponymy relation, we can identify: (a) pairs in which the response is the superordinate of the stimulus verb (such as *scrutare/vedere* ‘peer/see’); this relation has been identified in 22.8% of verb responses; (b) pairs in which the response is a subordinate of the stimulus verb (such as *parlare/chiacchierare* ‘talk/chat’); this relation is found among 5.9% of verb responses; (c) pairs of coordinate verbs. Two verbs (or better, verb senses) are called coordinate or also ‘sisters’ if they share the same superordinate (such as *correre/camminare* ‘run/walk’, both subordinates of a more generic MOTION verb such as *andare* ‘go’); this relation has been found in 11.7% of cases with respect to verb responses. Globally, troponymy relation

is the relation linking 40.5% of verb associates to their responses. Although quite widespread in all the verb classes, it can be seen that some verb classes are more strongly affected by troponymy relation in their structure. SUPPORT verbs have the highest percentage of superordinates among responses: almost all the verbs in this class elicited the generic support verb *aiutare* 'help'. On the contrary, CREATION verbs elicited a great number of verbs' subordinates. For example, the verb *creare* 'create', elicited its superordinate *fare* 'do, make', 4 times and 9 subordinates, namely *inventare* 'invent' (4), *costruire* 'build' (4), *ideare* 'conceive' (3), *comporre* 'compose' (2), *dipingere* 'paint' (1), *formare* 'form' (1), *produrre* 'produce' (1), *scrivere* 'write' (1).

A more in-depth inspection of the troponymy relations provides some insights into target verb properties: for instance target verbs with a large percentage of subordinates as associates are rather high frequency verbs (and therefore, given the strict correlation between frequency and semantic lightness we mentioned above, they are also conceptually more general), such as *creare* 'create'. As a matter of fact, the proportion of associate responses captured by this kind of relation increases as a function of target verb frequency, Pearson  $r = 0.14$ . Target verbs with a large number of superordinates as associates tend to be, on the contrary, rather specific, such as *supportare* 'support', *soccorrere* 'give aid to', etc.

Verbs in PERCEPTION class elicited a great amount of superordinates as well as of subordinates and coordinates. An analysis of association data concerning these verbs led us to recognize a real 'tree structure' for this sub-area of verb lexicon, with four lexicalised taxonomic levels. Consider the taxonomy arising from the verb *percepire* 'perceive': *percepire* is the highest-level verb, acting as superordinate of all PERCEPTION verbs; the next lower level contains relatively few verbs, basically one verb for each sense: *vedere* 'see', *sentire* and *udire* both 'hear', *toccare* 'touch', *annusare* 'smell', *gustare* 'taste'. Each of these verbs, as stimuli, elicited many subordinates: for example, *sentire* elicited *ascoltare* 'listen', 'hear with intention'; *toccare* 'touch', elicited *tastare* 'finger' and *sfiorare* 'barely touch'; *vedere* elicited *osservare* 'watch' and *guardare* 'look at'. This level is what might be called the 'bulge' for PERCEPTION taxonomies, that is to say, a level with far more lexicalized verbs than the other levels in the same hierarchy. This resembles what has been called the 'basic-level' in nominal hierarchies (the notion of a basic level within a hierarchical category structure was initially developed within the object domain: see Rosch et al. 1976 and Rosch 1978; it has been later applied to a wide variety of nonobjects domains, including actions and events: see

Morris & Murphy 1990). The lower level has few members: among the associates, we can find *scrutare* 'peer' and *intravedere* 'catch a glimpse of', *scorgere* 'catch sight of'. Associations seem to be strong between the level including the generic verbs for each sense (*vedere*, *sentire*, *toccare*, *annusare*, *gustare*) and their direct subordinates, the so-called basic level verbs: for example, *vedere* 'see', elicited *guardare* 'look at' (7 times) and *osservare* 'watch' (5 times). With a minor frequency, it elicited also its superordinate *percepire* 'perceive', (1), its indirect subordinate *scrutare* 'peer' (1) and its coordinate *sentire* 'hear' (1). Concerning the stimuli belonging to the basic level, these verbs elicited mostly their direct superordinates (whereas the higher-level verb *percepire* 'perceive' is almost never produced) and coordinates and, to a lesser extent, subordinates. Considering the basic level verb of seeing *guardare* 'look at', it evoked respectively *vedere* 'see' (13), *osservare* 'watch' (8), *scrutare* 'peer' (1). As expected, the semantic more elaborate troponyms, the lowest level verbs, tend to evoke their superordinates: *scrutare* 'peer', elicited *guardare* 'look at' (11), *osservare* 'watch' (8), *vedere* 'see' (2).

On the contrary, associations provided for some verb classes show a relatively flat structure: for example, CHANGE verbs and ASPECT verbs are linked almost totally to synonyms and antonyms. Virtually no other relation holds these verbs together. Thus, the organisation of this sub-area of the lexicon is flat rather than hierarchical: there are no superordinates and virtually no subordinates, so that CHANGE and ASPECT verbs seem to have a structure resembling rather that of adjectives (for the organisation of adjectives, see K.J. Miller 1998).

An interesting piece of information is provided by the verb-verb pairs for which we do not find a proper relationship among the relationships encoded in the WordNet architecture. As we can see in Table 7, they constitute a considerable percentage of verb associates (6%). First, it is worth noticing that among association pairs there are words that are linked not by semantic relations, but rather by various forms of collocational or idiomatic patterns. Collocational responses occur both among verb associates (an example being *dire* 'say', and the elicited response *fare* 'do, make', probably due to the common saying *tra il dire e il fare c'è di mezzo il mare* 'easier said than done') and among non-verbal associates (an example being *aiutare* 'help', and the elicited noun response *mano* 'hand', probably due to the idiomatic expression *dare una mano* 'help'). However, the vast majority of the association pairs we considered as related by an unknown relation represent instances of verb-verb relations not targeted by WordNet and ItalWordNet. For example, *entrare* 'enter, get in', was associated with the temporally pre-

ceding *aprire* 'open', and *cuocere* 'cook' (transform and make suitable for consumption by heating), with the temporally following *mangiare* 'eat'. In these pairs, the various relations regarding the temporal order of the subsumed events or activities are paired with the relation connecting an action or activity with the goal or purpose for which the action/activity is performed (it is assumed that one cooks something in order to eat it, or that one opens a door, in order to enter the room etc). Other instances of these purpose or goal relation are found among our experiment pairs, as we can see for example in the stimulus-response pairs *interrogare* 'conduct an examination or an interrogatory'/ *valutare* 'evaluate', *registrare* 'record'/ *ricordare* 'remember', or *leggere* 'read'/ *imparare* 'learn'. Moreover, we can find instances of events prototypically related to a common agent or part of a common cognitive frame (Fillmore 1982; Minsky 1975). For instance, the verb *interrogare* 'conduct an examination, test' elicited *far lezione* 'give a lecture', both activities typically connected with the role of teachers, and both parts of the sequence of activities that constitute the teaching script.

These examples are instantiations of verb relations not encoded in WordNet. For those cases, we think that our empirical association data provide a useful basis for evaluating the psycholinguistic salience of other non-classical relations, which could be eventually used to enhance the available lexical semantic networks.

#### *4.3. Semantic roles of nominal associates*

In a third phase, we have investigated the semantic roles realised by nominal associates elicited by the verb stimuli. As stated before, these associates constitute the majority of responses, 54.4% of the whole set of associates. Is it possible to recognise some patterns in the distribution of nominal responses across the set of experiment verbs? Guided by this question, we manually labelled the kind of semantic role that is realised by each nominal associate with respect to the target verb, that is, we tried to assign to each noun the proper semantic role that the concept denoted by the noun plays with respect to the action or state expressed by the verb. The list of semantic roles taken into account for the purposes of such a classification is provided in Appendix B.

In a following step, we summed the association frequencies with respect to a specific relationship, e.g., for the patients of the verb *scrivere* 'write', we summed over the frequencies of the various generated patients, i.e., *lettera* 'mail', *libro* 'book', *poesia* 'poem', *testo* 'text', *canzone* 'lyric' etc. Thus, we obtained a frequency distribution over semantic roles for each target verb. For instance, the most prominent

semantic roles instantiated by nominal associates for the object-drop verb *scrivere* are the patient (51% of associates) and the instrument (18.4% of the associates).

The most pervasive semantic role among the noun associates is the patient. Patients are provided by participants in 35.1% of the nominal associates. Speakers have produced patients as associations for almost all transitive verbs, although at varying degrees. For example, CREATION verbs like *costruire* 'build', or CONSUMPTION verbs like *mangiare* 'eat', elicited patients significantly more often than ELIMINATION verbs such as *rompere* 'break', or *uccidere* 'kill', even sharing exactly the same argumental structure.

Speakers produced patients as associations to stimulus verbs much more than agents (6.7% of nominal associates) and experiencers (1.9%). A look at the association data led us to conclude that agents and experiencers are usually produced as association only when the stimulus verb strongly implies in its meaning a particular kind of agent/experiencer. In these cases the verb lexically constrains the type of agent. For instance, whereas in the meaning of the verb *andare* 'go', there is no inherent reference to a particular class of involved agents (because, in fact, many kinds of entities can go), the verb *marciare* 'march', clearly involves the class of agents *soldati* ('soldiers'), *truppe* ('troops'), *esercito* ('army'), and the consequence is that agents constitute 55% of total associations for this verb.

The second most frequent role found among the association pairs is the relation that links a stimulus verb to the instrument used to perform the action denoted by that verb (7.6%). A look at the experiment data shows that the verbs that elicited the major number of nouns labeled as instruments are MOTION USING A VEHICLE verbs, with 33% of instrument responses followed by PERCEPTION verbs (among these verbs, 20% of responses denote a kind of instrument). If we try to compare words denoting instruments in the two cases, we can notice that the label 'instrument' applies to a vast range of 'objects'. In the former case instrument associates denote all vehicles and, more in general, artifacts. In the latter case, instrument associates are divided between nouns denoting the senses (*vista* 'sight', *tatto* 'touch', etc.) and nouns denoting the body parts involved in perception (the sense organs), with a large predominance of body parts on senses (that is, a verb like *vedere* 'see', elicited both the association *occhi* 'eyes', and the association *vista* 'sight', but the former was produced by 18 participants and the latter by 3 participants; this is not surprising if we consider the different degree of abstractness of nouns denoting body parts and nouns denoting senses).

Nouns denoting recipients are few, even considering only the subset of verbs which require three semantic arguments, such as TRANSFER OF POSSESSION or COMMUNICATION verbs. The same can be said for nouns playing the accompaniment role (e.g., *uscire/amici* 'go out/ friends'). On the contrary, the associates denoting the location in which an event/state takes place (5.6%) are quite numerous. In particular, it can be noticed that locations, together with goals and sources, are frequent associations for MOTION verbs (e.g., *correre/parco* 'run/park', *nuotare/piscina* 'swim/swimming pool', *volare/cielo* 'fly/sky') and for BE IN POSITION verbs (e.g., *stare/casa* 'stay/home'). The time in which the action or state denoted by the verbs takes place is specified in less than 1% of nominal associates (e.g., *nevicare/inverno* 'snow/winter').

Results (6.8%) are very frequent as responses to EMOTION verbs (e.g., *deludere/tristezza* 'disappoint/sadness', *preoccupare/ansia* 'worry/anxiety'). Concerning nouns denoting the possible causes for the action/state denoted by the stimulus verb (5.7%), it should be noticed that there are certain sub-areas of verb lexicon that are almost completely organised by this relation: some verbs denoting bodily processes are linked almost totally to nouns (and verbs) denoting the causes of these processes, as we can see from the associations provided for the verb *tremare* 'tremble': *freddo* 'cold', *paura* 'fear', *brividi* 'shivers', *emozione* 'emotion', *febbre* 'fever', *vento* 'wind'. In fact, they all denote the possible causes for which someone trembles (causal relation covers the 81% of the total associates for this verb). Similar results are found for verbs like *sbadigliare* 'yawn' (79%), *arrossire* 'blush' (67%), *starnutire* 'sneeze' (65%), *grattarsi* 'scratch' (62%). The percentage of causes among the whole set of associations for the class is 36.1%. To a lesser extent (16.2%), EMOTION verbs also elicited a quite massive number of causes, as we can see from the list of associations to the verb *struggersi* 'pine': indeed, causes (*dolore* 'pain', 14 times; *amore* 'love', 3; *rimpianto* 'regret', 2; *malinconia* 'gloom', 2) constitute 64% of total associations.

Altogether, although the classical set of semantic roles seems to be quite satisfactory in dealing with experiment responses, there remain associations whose link to the stimulus verb cannot properly be labelled with any of the aforementioned relations. For instance, we can consider the pairs *piovere/ombrello* 'rain/umbrella' (7 associations), or *insegnare/lavagna* 'teach/blackboard' (3), *interrogare/esame* 'test/exam' (3), *vestire/moda* 'wear/fashion' (3), *ballare/musica* 'dance/music' (6). A more suitable model for these associations could be provided by the notion of 'semantic frame', defined as "any system of

concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits” (Fillmore 1982: 111). In fact, the associates above fulfil frame-related roles: *rain* and *umbrella* express concepts that, although not related by structural semantic relations, are nonetheless related by ordinary human experience. In general, 6.2% of total responses can be interpreted as frame-related roles, and this type of relation appears particularly significant among WEATHER verbs and among TEACHING AND LEARNING verbs. Therefore, Frame Semantics might offer a natural account for a number of problematic phenomena that can not be captured by the core set of semantic roles that are traditionally adopted in linguistic theory.

### 5. Concluding remarks

This paper provided a detailed breakdown of the types of semantic relations and functional properties that are evoked by Italian verbs during a word association elicitation task.

The advantages of using a word association test as a tool to investigate semantic representations are numerous. Simplicity of data acquisition and great variety of semantic information extracted are only the major ones. Apart from these, another important advantage is that word association tests can give a relatively transparent measure of the information that is normally accessed when a word is heard or read as a result of lexical access. We assumed that the associations reflect highly salient linguistic and conceptual features of the verbs, and that, by collecting associations from multiple native speakers, we can gain a fine-grained measure of featural salience. In fact, a feature’s relative contribution to a word’s meaning can be weighted according to the number of speakers who produced that word. We also assumed that associations provided by speakers cover a wide range of possible semantic relations between words, and that they can provide useful empirical evidence to define the proper set of relations relevant to model the organization of the semantic lexicon.

For the approx. 5,700 verb associates we analysed, we identified classical WordNet relations for the vast majority of stimulus-response pairs (94%) and we discussed the distribution of these relations across the verb classes, addressing in this way the question of whether the same types of relations are salient for different types of verb classes. Furthermore, we tried to identify the remaining 6% of non-classical relations, such as temporal order, purpose, frame-relations, etc.

For the approx. 8,000 noun associates, we investigated the kinds of semantic roles that are realised by noun associates with respect to the target verb. Their overall distribution shed light on the fact that the noun associates are absolutely not restricted to those related to the argument structure of the target verbs. Less than a half of noun responses can behave as frame-slot fillers. The majority of nouns do not represent strictly subcategorized arguments. For instance, adjuncts such as the instruments used to perform the action denoted by the verb, the possible causes that produce the event, its results etc. are identified by the participants as salient features in the representation of verb's meaning as well. In Section 4.1, we mentioned the fact that the verb classes that evoked the highest percentage of noun responses are WEATHER verbs and BODILY PROCESSES verbs. This finding is quite surprising, since these are the classes whose verbs exhibit the minimum number of arguments (verbs like *piovere* 'rain', *nevicare* 'snow', *albeggiare* 'dawn' are zero-argument verbs; other weather verbs have one argument, such as *splendere* 'shine', *tramontare* 'set', whereas bodily processes verbs are mostly one-argument verbs, such as *dormire* 'sleep', *piangere* 'cry'). Our analysis about the semantic roles realised by noun associates can now explain clearly this finding: the percentage of noun responses turned out to be completely disconnected from the number of semantic arguments exhibited by the target verbs. For instance, BODILY PROCESSES verbs are particularly rich in nominal responses because they evoked, with a frequency higher than other classes, nouns denoting the possible causes as well as the possible results of the processes denoted by the verbs.

Data coming from association experiments can have important consequences with respect to distributional models of verb semantics. In data-intensive lexical semantics, words are commonly modeled by distributional vectors, and the relatedness of words is measured by vector similarity (Sahlgren 2006). The intuition underlying these approaches is that the meaning of a word is related to the distribution of words around. Crucial in distributional descriptions of word meaning is the choice of features to encode in vectorial representation. These features can be varied in nature: words co-occurring in a document, in a context window, or with respect to a word-word relationship, such as syntactic structure, syntactic and semantic valency, etc. Most previous work on distributional similarity has either focused on a specific word-word relation (such as Pereira et al. 1993 referring to a direct object noun for describing verbs), or used any dependency relation detected by the chunker or parser (e.g. Lin 1998). Our findings provide a further confirmation to what has been claimed by

Schulte im Walde & Melinger (2005; *to appear*): i.e., in order to encode the most of nominal information identified as somehow salient by the speakers, a representation based only on syntactically-based relations could be not completely adequate. These behavioral results support therefore the integration of window-base approaches into syntactically-based approaches.

Our findings may also have interesting consequences for modeling the semantic lexicon both in linguistics and in cognitive science. For instance, the high number of noun associates casts doubt on lexical architectures in which verbs are only (or mainly) organized in terms of verb-to-verb relations, such as for instance WordNet. Actually, an interesting element of novelty of ItalWordNet with respect to its American archetype is exactly represented by the fact that it also includes cross-PoS relations linking verbs to nouns. The salience of verb associates referring to different types of event participants also supports the importance of contextual setting information in conceptual representations. Although the relational nature of events is well-known, it is worth emphasizing that elicited associates do not only refer to the core set of verb arguments, but they also extend to prototypical causes, instruments, locations, etc. Indeed, we can claim that verb associations relate to the broad “scenario” and contextual setting of the event expressed by the verb. Our results are therefore consistent with the view by Barsalou (2005: 622) according to which “conceptual representations are contextualized dynamically to support diverse courses of goal pursuit”. The distribution of associate types can therefore be explained by assuming that when participants produce associations in a verb association task they access a highly contextualized representation of the event or situation expressed by the stimulus, possibly through its “virtual re-enactment” in a typical setting and with typical participants.

*Address of the Authors:*

Annamaria Guida, Center for Mind/Brain Sciences, University of Trento <annamaria.guida@unitn.it>

Alessandro Lenci, Department of Linguistics “T. Bolelli”, University of Pisa <alessandro.lenci@ilc.cnr.it>

## Notes

\* The authors would like to thank Sabine Schulte im Walde for her detailed advice and kind support in methodological and technical issues and Pier Marco Bertinetto for his valuable comments on the previous version of this work. We are also grateful to the anonymous referees for their useful suggestions.

<sup>1</sup> The main advantages and disadvantages of web experimenting are widely discussed in Reips (2002). The set of standards for web experimenting defined there was used in this work in order to use as appropriately as possible this technique.

<sup>2</sup> All of the analyses reported in this paper are based on response tokens; the type analyses show the same overall picture, apart from some cases that are explicitly pointed through the paper.

<sup>3</sup> Note that *impallidire* 'turn pale', is a change of state verb that affects the body, and for this reason could be classified both in BODILY PROCESSES class and in CHANGE class, depending on which meaning component one prefers to bring into focus.

<sup>4</sup> Quite interestingly, the far most frequent associate for *impallidire* is not its root adjective *pallido* (provided by 2 participants) but another adjective, namely *bianco*, 'white' (provided by 16 participants), bearing exactly the same semantic relation to the stimulus. An interesting question is why *pallido* is not immediately evoked by the verb *impallidire*. We did not select the experiment verbs in order to investigate eventual processes of morphological decomposition. However, the results provided by *impallidire* and the other deadjectival verbs, besides the associations evoked by the verb *ricominciare* 'start again' (derived from the verb *cominciare* 'start', by adding the iterative prefix *ri-*) seem to point out that a morphological decomposition does not operate at this level. *Ricominciare* evoked *cominciare* only once but its synonym *iniziare* 7 times. Therefore, in those cases in which two words bearing exactly the same semantic relation to the stimulus were available to the speakers – one being morphologically related to the stimulus (more precisely, one being the word from which the stimulus is morphologically derived) and the other being not – the speakers produced the morphological related word with much lower frequency than the other. We are aware that we are dealing with a too limited number of derived target verbs to be able to formulate hypotheses with respect to the morphological processes within word association task, which could be a concern of future work.

<sup>5</sup> Troponymy is a hierarchical relation between two events such that the former represents a manner elaboration of the latter; e.g., *craw* is a troponym of *move* because crawling is a way of moving.

<sup>6</sup> The number of relations in WordNet was kept deliberately small, and lumping together several subrelations as well as ignoring certain semantic distinctions seemed justified for several reasons. For a discussion about this choice, see Fellbaum (ed.) (1998). In this paper we use these semantic relations labels according to the definitions by Miller et al. (1990) and Fellbaum (1998). As a consequence, in this paper, like in WordNet, the relation of backward presupposition (holding between verbs such as *trovare* 'search' and *cercare* 'find') is subsumed under lexical entailment and antonymy covers also those pairs traditionally called converses (like *vendere* 'sell' and *comprare* 'buy').

*Bibliographical References*

- BARONI Marco et al. 2004. Introducing the *La Repubblica* corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. In LINO Maria Teresa et al. (eds.). *Proceedings of LREC 2004* (26-28 May 2004). Lisbon, Portugal. 1171-1174.
- BARSALOU Lawrence W. 2005. Situated conceptualization. In Henry COHEN & Claire LEFEBVRE (eds.). *Handbook of categorization in cognitive science*. St. Louis: Elsevier. 619-650.
- BERTINETTO Pier Marco 1986. *Tempo, Aspetto e Azione nel verbo italiano. Il sistema dell'indicativo*. Firenze: Accademia della Crusca.
- BERTINETTO Pier Marco & Mario SQUARTINI 1995. An attempt at defining the class of gradual completion verbs. In BERTINETTO Pier Marco et al. *Temporal reference, aspect and actionality*. Vol I: *Semantic and syntactic perspectives*. Torino: Rosenberg & Sellier. 11-27.
- BURGESS Curt 1998. From simple associations to the building blocks of language: Modeling meaning in memory with the Hal model. *Behavior Research Methods, Instruments & Computers* 30. 188-198.
- CLARK Herbert H. 1971. Word associations and linguistic theory. In Lyons John (ed.). *New Horizons in Linguistics*. Harmondsworth: Penguin. 271-286.
- FELLBAUM Christiane 1998. *A semantic network of English verbs*. In FELLBAUM Christiane (ed.) 1998. 69-104.
- FELLBAUM Christiane (ed.) 1998. *WordNet – An Electronic Lexical Database*. Cambridge, MA: MIT Press.
- FERNÁNDEZ Ana et al. 2004. Free association norms for the Spanish names of the Snodgrass & Vanderwart pictures. *Behavior Research Methods, Instruments & Computers* 36(3). 577-583.
- FERRAND Ludovic & F.-Xavier ALARIO 1998. French word association norms for 366 names of objects. *L'Annee Psychologique* 98(4). 659-709.
- FILLMORE Charles J. 1968. The case for case. In BACH Emmon & Robert T. HARMS (eds.). *Universals in linguistic theory*. New York, NY: Holt, Rinehart and Winston. 1-90.
- FILLMORE Charles J. 1982. Frame semantics. In Linguistic Society of Korea (ed.). *Linguistics in the morning calm*. Seoul: Hanshin Publishing Co. 111-137.
- GALTON Francis 1880. Psychometric experiments. *Brain* 2. 149-162.
- GLENBERG Arthur M. & Michael P. KASCHAK 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9. 558-569.
- HERINGER Hans J. (1986). The verb and its semantic power: Association as the basis for valence. *Journal of Semantics* 4. 79-99
- JACKENDOFF Ray 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- JACKENDOFF Ray 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- KENT Grace H. & Aaron J. ROSANOFF 1910. A study of association in insanity. *American Journal of Insanity* 67 (37-96). 317-390.
- KISS George R. et al. 1973. An associative thesaurus of English and its computer analysis. In A.J. AITKEN, R.W. BAILEY & N. HAMILTON-SMITH (eds.).

- The Computer & Literary Studies*. Edinburgh: Edinburgh University Press. URL <http://www.eat.rl.ac.uk/>.
- LAUTESLAGER Max, Theo SCHAAP & Dick SCHIEVELS 1986. *Schriftelijke woordassociatienormen voor 549 Nederlandse zelfstandige naamwoorden*. Lisse: Swets & Zeitlinger.
- LEVIN Beth 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: The University of Chicago Press.
- LEVIN Beth & Malka RAPPAPORT HOVAV 1992. The lexical semantics of verbs of motion: the perspective from unaccusativity. In ROCA Iggy (ed.). *Semantic structure: its role in grammar*. Berlin: Mouton de Gruyter.
- LIN Dekang 1998. Automatic retrieval and clustering of similar words. *Proceedings of the 17th International Conference on Computational Linguistics* (10-14 August). Montreal, Canada. 768-774.
- McEVOY Cathy L. & Douglas L. NELSON 1982. Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology*, 95. 581-634.
- McKOOON Gail & Roger RATCLIFF 1992. Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18. 1155-1172.
- McNAMARA Timothy P. 2005. *Semantic priming: Perspectives from memory and word recognition*. New York, NY: Psychology Press.
- MCRÆE Ken & Stephen BOISVERT 1998. Automatic Semantic Similarity Priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(3). 558-72.
- MCRÆE Ken et al. 2005. Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods* 37. 547-559.
- MELINGER Alissa & Andrea WEBER 2006. *Database of nouns associations for German*. URL: [www.coli.uni-saarland.de/projects/nag/](http://www.coli.uni-saarland.de/projects/nag/).
- MELINGER Alissa, Sabine SCHULTE IM WALDE & Andrea WEBER 2006. Characterizing Response Types and Revealing Noun Ambiguity in German Association Norms. *Proceedings of the EACL Workshop 'Making Sense of Sense: Bringing Computational Linguistics and Psycholinguistics together'* (4 April 2006). Trento, Italy. 41-48.
- MILLER George A. et al. 1990. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography* 3 (4). 235-244.
- MILLER Katherine J. 1998. Modifiers in WordNet. In FELLBAUM Christiane (ed.) 1998. 47-68.
- MINSKY Marvin 1975. A framework for representing knowledge. In WINSTON Patrick Henry (ed.). *The psychology of computer vision*. New York, NY: McGraw-Hill. 211-277.
- MORRIS M.W. & Gregory MURPHY 1990. Converging operations on a basic level in event taxonomies. *Memory & Cognition* 18. 407-418.
- NELSON Douglas L., David J. BENNETT & Todd W. LEIBERT 1997. One step is not enough: Making better use of association norms to predict cued recall. *Memory & Cognition* 25. 785-796.
- NELSON Douglas L., Cathy L. McEVOY & Thomas A. SCHREIBER 1998. *The University of South Florida word association, rhyme, and word fragment norms*. URL: <http://www.usf.edu/FreeAssociation/>.

- NELSON Douglas L., Cathy L. McEVOY & Simon DENNIS 2000. What is free association and what does it measure? *Memory & Cognition* 28. 887-899.
- NELSON Douglas L. & Nan ZHANG 2000. The ties that bind what is known to the recall of what is new. *Psychonomic Bulletin & Review* 7 (4). 604-617.
- PALERMO D.S. & J.J. JENKINS 1964. *Word Association Norms: Grade School through College*. Minneapolis, MN: University of Minnesota Press.
- PEREIRA Fernando, Naftali TISHBY & Lillian LEE 1993. Distributional clustering of English words. *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics* (22-26 June 1993). Columbus, OH. United States. 183-190.
- PERESSOTTI Francesca, Francesca PESCIARELLI & Remo JOB 2002. Le associazioni verbali PD-DPSS: norme per 294 parole. *Giornale italiano di psicologia* 29. 153-170.
- PLAUT David C. 1995. Semantic and associative priming in a distributed attractor network. In MOORE Johanna D. & Jill F. LEHMAN (eds.). *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (August 2003), Pittsburg, PA. Vol 17. 37-42.
- REIPS Ulf-Dietrich 2002. Standards for Internet-based experimenting. In *Experimental Psychology* 49 (4). 243-256.
- ROSCH Eleanor 1978. Principles of categorization. In ROSCH Eleanor & Barbara B. LOYD (eds.). *Cognition and Categorization*. Hillsdale, NJ: Erlbaum. 27-48.
- ROSCH Eleanor et al. 1976. Basic objects in natural categories. *Cognitive Psychology* 8. 382-439.
- ROVENTINI Adriana et al. 2000. ItalWordNet: a Large Semantic Database for Italian. In Proceedings of LREC 2000 (31 May-2 June 2000). Athens, Greece. Vol. II, 783-790.
- RUSSELL Wallace A. 1970. The complete German language norms for responses to 100 words from the Kent-Rosanoff word association test. In POSTMAN Leo & Geoffrey KEPPEL (eds.). *Norms of Word Association*. New York, NY: Academic Press. 53-94.
- RUSSELL Wallace A. & O.R. MESECK 1959. Der Einfluss der Assoziation auf das Erinnern von Worten in der deutschen, französischen und englischen Sprache. *Zeitschrift für Experimentelle und Angewandte Psychologie* 6. 191-211.
- SAHLGREN Magnus 2006. *The Word-Space Model: Using distributional analysis to represent syntagmatic and paradigmatic relations between words in highdimensional vector spaces*. Stockholm University: Department of Linguistics. PhD dissertation.
- SCHULTE IM WALDE Sabine 2006(a). Can Human Associations Help Identify Salient Features for Semantic Verb Classification? In *Proceedings of the 10th Conference on Computational Natural Language Learning* (8-9 June, 2006). New York, NY. 69-76.
- SCHULTE IM WALDE Sabine 2006(b). Experiments on the Automatic Induction of German Semantic Verb Classes. *Computational Linguistics* 32 (2). 159-194.

- SCHULTE IM WALDE Sabine 2008. Human associations and the choice of features for semantic verb classification. *Research on Language and Computation* 6(1). 79-111.
- SCHULTE IM WALDE Sabine & Alissa MELINGER 2005. Identifying Semantic Relations and Functional Properties of Human Verb Associations. *Proceedings of the joint Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (6-8 October 2005). Vancouver, Canada. 612-9.
- SCHULTE IM WALDE Sabine & Alissa MELINGER *to appear*. An in-depth look into the co-occurrence distribution of semantic associates. In LENCI Alessandro (ed.). *Italian Journal of Linguistics*. Special Issue on *From Context to Meaning: Distributional Models of the Lexicon in Linguistics and Cognitive Science*.
- SCHULTE IM WALDE Sabine et al. *submitted*. An empirical characterisation of response types in German association norms. *Research on Language and Computation*.
- SLOBIN Dan 1996. Two ways of travel: verbs of motion in English and Spanish. In SHIBATANI Masayoshi & Sandra A. THOMPSON (eds.). *Grammatical constructions: their form and meaning*. Oxford: Clarendon Press.
- SPENCE Donald P. & Kimberly C. OWENS 1990. Lexical co-occurrence and association strength. *Journal of Psycholinguistic Research* 19. 317-330.
- TALMY Leonard 1985. Lexicalization patterns: Semantic structure in lexical forms. In SHOPEN Timothy (ed.). *Language typology and syntactic description*. Vol. 3: *Grammatical categories and the lexicon*. Cambridge: Cambridge University Press. 57-149.
- TANENHAUS Micheal K., Greg N. CARLSON & John C. TRUESWELL 1989. The role of thematic structure in interpretation and parsing. *Language and Cognitive Processes* 4 (3-4). 1211-1234.
- VENDLER Zeno 1967. Verbs and Times. In VENDLER Zeno. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press. 97-121.
- WU Ling & Lawrence W. BARSALOU *submitted*. *Grounding concepts in perceptual simulation: I. Evidence from property generation*.

## *Appendices*

### *A. Experiment verbs and classes*

Appendix A presents a full list of the 312 verbs included in the study and their classification into semantic classes. We defined a taxonomy with 18 semantic classes. The following table lists all the classes, accompanied by their sub-classes and the respective verbs. The Italian verbs are provided with a coarse translation into English, given in brackets. In the case of polysemous verbs, we translated only the sense consistent with the semantic class assigned to the verb.

Verb classes	Verbs	
COGNITION	<i>sapere</i> (know), <i>pensare</i> (think), <i>decidere</i> (decide), <i>ricordare</i> (remember), <i>credere</i> (believe), <i>sostenere</i> (believe), <i>conoscere</i> (know), <i>considerare</i> (consider), <i>ritenere</i> (deem), <i>dimenticare</i> (forget), <i>giudicare</i> (judge), <i>immaginare</i> (imagine), <i>ignorare</i> (ignore), <i>ipotizzare</i> (hypothesize), <i>sospettare</i> (suspect), <i>dubitare</i> (doubt), <i>supporre</i> (suppose), <i>indovinare</i> (guess), <i>fantasticare</i> (fantasize)	
DESIRE	WISH	<i>volere</i> (want), <i>cercare</i> (look for), <i>desiderare</i> (wish), <i>aspirare</i> (aspire), <i>ambire</i> (hanker), <i>bramare</i> (long)
	NEED	<i>occorrere</i> (be required), <i>richiedere</i> (require), <i>necessitare</i> (need)
Emotion	<i>preoccupare</i> (worry), <i>gridare</i> (shout), <i>sorridere</i> (smile), <i>ridere</i> (laugh), <i>divertire</i> (amuse), <i>urlare</i> (shout), <i>deludere</i> (disappoint), <i>spaventare</i> (scare), <i>arrabbiarsi</i> (get angry), <i>infuriarsi</i> (flare up), <i>rallegrare</i> (gladden), <i>annoiare</i> (bore), <i>impaurire</i> (frighten), <i>amareggiare</i> (sadden), <i>gioire</i> (rejoice), <i>addolorare</i> (pain), <i>struggersi</i> (pine)	
PERCEPTION	<i>vedere</i> (see), <i>sentire</i> (feel, hear), <i>guardare</i> (look at), <i>toccare</i> (touch), <i>ascoltare</i> (listen), <i>percepire</i> (perceive), <i>intravedere</i> (catch a glimpse of), <i>scorgere</i> (catch sight of), <i>accarezzare</i> (caress), <i>scrutare</i> (peer), <i>fiutare</i> (smell), <i>gustare</i> (taste), <i>annusare</i> (sniff), <i>assaggiare</i> (sample), <i>tastare</i> (touch, feel)	
COMMUNICATION	<i>dire</i> (say, tell), <i>parlare</i> (talk), <i>chiedere</i> (ask), <i>scrivere</i> (write), <i>rispondere</i> (answer, reply), <i>raccontare</i> (tell, narrate), <i>annunciare</i> (announce), <i>affermare</i> (assert), <i>leggere</i> (read), <i>proporre</i> (propose), <i>ripetere</i> (repeat), <i>riferire</i> (report), <i>discutere</i> (discuss, argue), <i>negare</i> (deny), <i>citare</i> (quote), <i>segnalare</i> (signal), <i>interrogare</i> (question, interrogate, test), <i>comunicare</i> (communicate), <i>telefonare</i> (phone), <i>scherzare</i> (joke), <i>domandare</i> (ask), <i>dettare</i> (dictate), <i>pregare</i> (beg), <i>ribattere</i> (reply, refute), <i>litigare</i> (quarrel, argue), <i>negoziare</i> (negotiate), <i>dialogare</i> (converse), <i>avvisare</i> (inform, warn), <i>insinuare</i> (insinuate, hint at), <i>chiacchierare</i> (chat), <i>faxare</i> (fax)	
TEACHING AND LEARNING	TEACHING	<i>spiegare</i> (explain), <i>esporre</i> (expound), <i>descrivere</i> (describe), <i>insegnare</i> (teach), <i>illustrare</i> (illustrate)
	LEARNING	<i>studiare</i> (study), <i>imparare</i> (learn), <i>apprendere</i> (learn), <i>memorizzare</i> (memorize)

*Semantic properties of word associations to Italian verbs*

VERBS INVOLVING THE BODY	DRESSING:	<i>vestire</i> (dress, wear), <i>indossare</i> (wear, put on), <i>mettersi</i> (put on), <i>spogliare</i> (undress), <i>svestire</i> (undress)
	BODILY PROCESSES	<i>piangere</i> (cry), <i>dormire</i> (sleep), <i>respirare</i> (breathe), <i>tremare</i> (tremble), <i>addormentarsi</i> (fall asleep), <i>svenire</i> (faint), <i>grattarsi</i> (scratch), <i>impallidire</i> (turn pale), <i>arrossire</i> (blush), <i>singhiozzare</i> (sob), <i>ansimare</i> (pant), <i>pettinare</i> (comb), <i>sbadigliare</i> (yawn), <i>tossire</i> (cough), <i>russare</i> (snore), <i>starnutire</i> (sneeze)
	GESTURES/SIGNS INVOLVING BODY PARTS	<i>scuotere</i> (shake), <i>annuire</i> (nod), <i>applaudire</i> (clap), <i>fischiare</i> (whistle), <i>mimare</i> (mime)
	BODILY STATE AND DAMAGE TO THE BODY:	<i>soffrire</i> (suffer, be in pain), <i>ferire</i> (wound, hurt), <i>bruciare</i> (burn), <i>picchiare</i> (beat), <i>violentare</i> (rape), <i>patire</i> (suffer), <i>scottare</i> (burn), <i>maltrattare</i> (maltreat), <i>pungere</i> (prick, sting), <i>prudere</i> (itch), <i>morsicare</i> (bite)
MOTION	<i>andare</i> (go) [superordinate]	
	MANNER OF MOTION	<i>correre</i> (run), <i>saltare</i> (jump), <i>camminare</i> (walk), <i>ballare</i> (dance), <i>marciare</i> (march), <i>passaggiare</i> (stroll), <i>vagare</i> (roam), <i>nuotare</i> (swim), <i>rotolare</i> (roll), <i>barcollare</i> (stagger)
	MANNER OF MOTION USING A VEHICLE	<i>guidare</i> (drive), <i>volare</i> (fly), <i>sciare</i> (ski), <i>pedalare</i> (cycle), <i>remare</i> (row), <i>pattinare</i> (skate)
	DIRECTION	<i>venire</i> (come), <i>tornare</i> (come back), <i>entrare</i> (go into, get into), <i>uscire</i> (go out), <i>partire</i> (leave), <i>scendere</i> (go down, get off), <i>salire</i> (go up, get on), <i>cadere</i> (fall), <i>girare</i> (turn, go around), <i>proseguire</i> (go on), <i>avanzare</i> (to move forward), <i>fuggire</i> (run away, flee), <i>scivolare</i> (slide, slip), <i>schizzare</i> (splash, squirt), <i>svoltare</i> (turn)

POSITION	BRING INTO POSITION	<i>mettere</i> (put, place, set), <i>porre</i> (put, place, set), <i>togliere</i> (take away, take off), <i>introdurre</i> (insert), <i>spostare</i> (move), <i>sollevare</i> (raise, lift), <i>collocare</i> (place), <i>levare</i> (remove), <i>appendere</i> (hang), <i>riporre</i> (put, put back, put away), <i>posizionare</i> (set, position)
	BE IN POSITION	<i>stare</i> (stay), <i>esserci</i> (be there, be around), <i>trovarsi</i> (be, be situated), <i>restare</i> (remain), <i>rimanere</i> (remain), <i>vivere</i> (live), <i>occupare</i> (fill, take up, squat), <i>sedere</i> (sit), <i>abitare</i> (live, inhabit), <i>giacere</i> (lie)
TRANSFER OF POSSESSION	GIVING	<i>dare</i> (give), <i>portare</i> (bring, carry), <i>pagare</i> (pay), <i>offrire</i> (offer), <i>vendere</i> (sell), <i>fornire</i> (supply, provide), <i>mandare</i> (send), <i>trasferire</i> (transfer), <i>distribuire</i> (distribute, give out), <i>restituire</i> (return, give back), <i>regalare</i> (present), <i>prestare</i> (lend), <i>donare</i> (present), <i>spartire</i> (share out), <i>rimborsare</i> (refund), <i>porgere</i> (give, offer)
	OBTAINING	<i>trovare</i> (find), <i>prendere</i> (take), <i>ottenere</i> (obtain), <i>accettare</i> (accept), <i>ricevere</i> (receive), <i>comprare</i> (buy), <i>ereditare</i> (inherit), <i>affittare</i> (rent), <i>appropriarsi</i> (appropriate)
CREATION	<i>fare</i> (do, make), <i>creare</i> (create), <i>produrre</i> (produce), <i>costruire</i> (build), <i>generare</i> (generate), <i>partorire</i> (beget), <i>innalzare</i> (raise), <i>erigere</i> (erect)	
CHANGE OF STATE	<i>diventare</i> (become), <i>cambiare</i> (change), <i>nascere</i> (be born), <i>chiudere</i> (close), <i>morire</i> (die), <i>crescere</i> (grow), <i>aumentare</i> (increase), <i>migliorare</i> (improve), <i>modificare</i> (modify), <i>rinnovare</i> (renew), <i>allargare</i> (enlarge), <i>diminuire</i> (reduce, diminish), <i>aprire</i> (open), <i>allungare</i> (lengthen, stretch), <i>peggiorare</i> (get worse, make worse), <i>restaurare</i> (restore), <i>pulire</i> (clean), <i>riparare</i> (repair), <i>invecchiare</i> (grow old, age), <i>perfezionare</i> (perfect), <i>aggiustare</i> (repair), <i>restringere</i> (narrow), <i>sporcare</i> (dirty), <i>accorciare</i> (shorten), <i>dimagrire</i> (slim), <i>ingrassare</i> (fatten), <i>cuocere</i> (cook), <i>decrescere</i> (decrease), <i>rimpicciolire</i> (make smaller)	
CONSUMPTION AND ELIMINATION	CONSUMPTION	<i>mangiare</i> (eat), <i>consumare</i> (consume), <i>bere</i> (drink), <i>divorare</i> (devour), <i>logorare</i> (use up)
	ELIMINATION	<i>uccidere</i> (kill), <i>colpire</i> (hit), <i>cancellare</i> (erase, wipe out, rub out), <i>distruggere</i> (destroy), <i>rompere</i> (break), <i>eliminare</i> (eliminate), <i>abbattere</i> (pull down, knock down), <i>spezzare</i> (break, split), <i>spaccare</i> (break, split), <i>devastare</i> (ravage), <i>frantumare</i> (shatter, crash)

*Semantic properties of word associations to Italian verbs*

MEASURE	<i>registrare</i> (record), <i>contare</i> (count), <i>contenere</i> (contain), <i>costare</i> (cost), <i>pesare</i> (weigh), <i>misurare</i> (measure), <i>stimare</i> (value, estimate)
SUPPORT	<i>seguire</i> (follow, coach), <i>servire</i> (serve), <i>aiutare</i> (help), <i>salvare</i> (save), <i>soccorrere</i> (succour, assist), <i>supportare</i> (support)
WEATHER	<i>piovare</i> (rain), <i>congelare</i> (freeze), <i>soffiare</i> (blow), <i>tuonare</i> (thunder), <i>tramontare</i> (set), <i>nevicare</i> (snow), <i>splendere</i> (shine), <i>grandinare</i> (hail), <i>imbrunire</i> (get dark), <i>diluviare</i> (pour, rain in torrents), <i>albeggiare</i> (dawn), <i>annuvolare</i> (cloud)
ASPECT	<i>continuare</i> (go on), <i>cominciare</i> (begin), <i>finire</i> (end), <i>iniziare</i> (begin), <i>ricominciare</i> (begin again), <i>terminare</i> (end), <i>cessare</i> (stop)
PERFORMATIVE VERBS	<i>costringere</i> (compel), <i>imporre</i> (impose), <i>condannare</i> (sentence), <i>promettere</i> (promise), <i>minacciare</i> (threaten), <i>eleggere</i> (elect), <i>nominare</i> (appoint), <i>ordinare</i> (order), <i>vietare</i> (forbid), <i>obbligare</i> (oblige), <i>comandare</i> (order), <i>intimare</i> (enjoin)

*B. Semantic roles for nominal associates*

Semantic roles were introduced in generative grammar during the mid-1960s and early 1970s (Fillmore 1968, Jackendoff 1972) as a way of classifying the arguments of natural language predicates into a closed set of participant types which were regarded to have a special status in grammar. Therefore, semantic roles attempt to capture similarities and differences in verb meanings that are reflected in argument expression and that contribute to the mapping from semantics to syntax. The literature records scores of proposals for sets of semantic roles. A list of the roles we used as a basis for the classification of our nominal associates is given below. We also provided some examples for each role from our data.

Semantic role	Definition	Examples from our data set
Agent	A participant doing or causing something (possibly intentionally) in the event	<i>assassino</i> ‘killer’, for <i>uccidere</i> ‘kill’; <i>soldati</i> ‘soldiers’, for <i>marchiare</i> ‘march’
Patient	A participant which the verb characterizes as having something happen to it, and as being affected by what happens to it (for example, changing its position or condition)	<i>cibo</i> ‘food’, for <i>mangiare</i> ‘eat’; <i>libro</i> ‘book’, for <i>scrivere</i> ‘write’
Experiencer	A participant who is characterized as aware of something or as experiencing some stimulus	<i>pubblico</i> ‘audience’, for <i>annoiare</i> ‘annoy’
Instrument	The instrument by which the event or situation denoted by the predicate is carried out	<i>penna</i> ‘pen’, for <i>scrivere</i> ‘write’
Recipient	An entity receiving something	<i>poveri</i> ‘poors’, for <i>distribuire</i> ‘distribute’

Time	The time in which the event or situation denoted by the predicate is situated	<i>inverno</i> 'winter', for <i>nevicare</i> 'snow'
Location	The place where the event or situation denoted by the predicate is situated	<i>piscina</i> 'swimming pool' for <i>nuotare</i> 'swim'
Source	The location or entity from which motion proceeds	<i>casa</i> 'house, home' for <i>uscire</i> 'go out'
Goal	The location or entity in the direction of which something moves	<i>casa</i> 'house, home' for <i>andare</i> 'go'
Accompaniment	An entity that participates in an event or situation in close association with an agent, causer, or affected entity	<i>amici</i> 'friends' for <i>uscire</i> 'go out'
Cause	An entity or state which specifies a possible cause determining the event or situation denoted by the predicate	<i>paura</i> 'fear' or <i>freddo</i> 'cold', for <i>tremare</i> 'tremble'
Result	An entity or state which specifies a possible result determined by the action or situation denoted by the predicate	<i>tristezza</i> 'sadness' for <i>deludere</i> 'disappoint'; <i>dolore</i> 'pain', for <i>cadere</i> 'fall'