

**RITA MARINELLI**  
**REMO BINDI**  
**ADRIANA ROVENTINI**  
 Istituto di Linguistica Computazionale  
 Pisa, Italy  
 rita.marinelli@ilc.cnr.it , remo.bindi@ilc.cnr.it, adriana.roventini@ilc.cnr.it

### ***Metonymic and Metaphorical Uses of Proper Names***

#### **Abstract**

In this paper we describe a research we are carrying out in the framework of ItalWordNet, a large lexical-semantic database containing semantic information for about 50,000 synsets of nouns, verbs, adjectives, adverbs, and a subset of proper names, which is continuously enriched and updated at the Institute for Computational Linguistics in Pisa.

The research concerns the proper names considered from a twofold aspect: their coding in the lexical semantic database and their concrete use, as it is evidenced within a large corpus of the contemporary written Italian language. In particular the semantic relations involving the proper names and their senses (literal, derived and extended) are taken into consideration.

Many proper names turn out to be the basis for many extensions of meaning, so in the paper we analyse concretely in the corpus a group of them showing many types of derivatives and sense extensions generated by means of lexical rules that operate as “generative factors” (Pustejovsky, 2001).

The corpus lines analysis confirms a rich regular polysemy regarding this category of nouns, so, for a few cases, we propose to represent this phenomenon by introducing specific semantic relations in the database.

#### **1. Introduction**

ItalWordNet (IWN) has been built enlarging the Italian WordNet developed in the framework of the European project EuroWordNet (EWN) by codifying new grammatical categories (adjectives and adverbs) and a subset of proper names (pn). We aim at focusing this subset, mainly to achieve a well reasoned and structured enlarging of the database, also through the deeper study of the semantic relations involving the pn, on the basis of the experience carried out in building the IWN database.

The most important relation, within IWN, is the ‘synonymy’ relation: it applies to the variants of a synset allowing to interchange the synonyms (or variants) in at least one proposition, and this kind of relation is valid also for the set of pn as formalized below:

$$a = b \Leftrightarrow \{a.f(a)\} = \{b.f(b)\}$$

In IWN, the relation ‘*belongs to class*’ and its reversed ‘*has instance*’ connect an individual (instance) with the class it belongs to (‘inherence’ propositions); the hyponymy relation ‘*is a*’ (class-class or relation propositions) is not applicable to pn.

The subset of pn consists, up to now, of more than 4000 pn, originated from a first subset of geographic pn, further increased with data coming from sources of several type: atlases, Web sites, lists of various kind.

More than 200 classes of pn were defined. The database was also enriched encoding other relations involving pn. In fact, it was evidenced that many pn are the basis of many substantives and adjectives as their derivatives (e.g.: Nicot-nicotine) and that, when an adjective derived from the pn does not exist, pn are used very often in appositive/attributive position, e.g.: Morse code. (Marinelli and Roventini, 2002).

The *pertains\_to* relation and its reverse *has\_pertained*, has been used both in WN and in EWN. They allow the link of a noun with a relational adjective. This relation applies either between synsets or between synsets and instances: it connects 2<sup>nd</sup> order entities with 1<sup>st</sup> order entities, or 2<sup>nd</sup> order entities and instances:

<b>Florence</b>	<b><i>has_pertained</i></b>	<b>Florentine</b>
<b>philosophical</b>	<b><i>pertains_to</i></b>	<b>philosophy</b>

Like the other grammatical categories, also pn were linked with WordNet 1.5 by means of equivalence relations. The *eq\_synonymy* is used to map pn with an equivalent instance in WN; the *eq\_belongs\_to\_class*, that was not present in EWN: it has been codified in IWN to link pn to the generic belonging class when they have no equivalent in WordNet.

In the following examples all the types of relations so far encoded for this subset are shown:

<b>Italia</b>	<b><i>belongs_to_class</i></b>	<b>nazione (nation)</b>
<b>italiano</b>	<b><i>pertains_to</i></b>	<b>Italia (Italy)</b>
<b>Italia</b>	<b><i>derivation</i></b>	<b>italianità (italian world)</b>
<b>Italia</b>	<b><i>eq_synonym</i></b>	<b>Italy</b>
<b>Moldova</b>	<b><i>eq_belongs_to_class</i></b>	<b>nation</b>

#### **2. Proper Names and Polysemy**

In this paper we would refer more widely about the relations involving the pn, in particular codifying the relation between the pn and the senses (literal, derived and extended). Regular polysemy has been widely studied, closely connected with linguistic phenomena such as metonymy.

Our purpose is to highlight particular cases of polysemy found in codifying pn.

The study of polysemy in pn may, in our advice, constitute a useful contribute for a more general ‘theory of polysemy’ concerning also the other grammatical categories. To study the sense shifting mechanisms in this subset is helpful for understanding and describing more sophisticated processes of transposition that are enriching the written and spoken language in every day life.

We consider pn as the basis for many extensions of meaning: this may happen when “a more general human metarepresentational capacity” is exploited (Papafragou, 1995). In fact, many types of derivatives and sense extensions are generated, by means of lexical rules that operate as “generative factors” (Pustejovsky, 1995). Novel usages of a word form can be derived through productive application of a lexical rule; therefore we want to represent these lexical rules codifying new semantic relations in the database.

A polysemic production happens by means of a metaphoric use of the pn (Your husband is a Croesus), by means of metonymy (to read Dante), by means of an antonomasia (the Aquinate), or by means of lexical mechanisms like analogy/synecdoche (he would like to drink a Bloody Mary).

Many cases of polysemy can be understood as the result of generative mechanisms. So we want to give prominence to the polysemy of pn to confirm the linguistic manifestation(s) of the faculty for generative categorization and compositional thought (Pustejovsky, 2001), and to highlight how “productive processes which require generative lexical mechanisms” (Copestake and Briscoe, 1996) also apply to pn.

In our experience, it has been noticed that some deviations from the literal reference are present regularly (Nunberg, 1996), when considering some particular belonging classes, sharing regular semantic relationships; hereafter some examples are shown:

- Town/citizens e.g.: Roma has now its new mayor
- Nation/people e.g.: Only the 60% of Italy voted
- Building/person/Institution e.g.: the Quirinale welcomes the princess
- Person/corporation e.g.: Lacoste, Ford, Skoda, Ferrari
- Place/product e.g.: Madras, Sheffield, Shetland
- Writer/literary work e.g.: I like Neruda
- Artist/work of art e.g.: a Picasso was stolen
- Craftsman/artifact e.g.: a Stradivari was found in an old chest
- Corporation/product e.g.: the Ferrari won the Formula 1
- University/town e.g.: Bologna is a good law faculty
- Place/battle/defeat or victory e.g.: Waterloo, Caporetto
- Physician/unit of measurement e.g.: Hertz, Baud, Ampere
- Scientist/discovery/medical analysis e.g.: Doppler
- Musician/composition e.g.: Brahams is too difficult for me
- Region/skiing technique e.g.: Telemark
- Athlete’s name/technique e.g.: Fosbury, Cassina
- Name of the grapes/wine e.g.: Bordeaux

Until now only the ‘*derivation*’ relation has been used e.g.:

**Hertz1 belongs to class fisico (physician)**  
**hertz2 belongs to class unità di misura (unit of measurement)**  
**Hertz1 derivation hertz2**

It connects variants belonging to different PoSs (Parts of Speech) and applies both to the 1<sup>st</sup> and to the 2<sup>nd</sup> order entities as shown in the examples below:

**Dalton derivation daltonico (daltonic)**  
**Pastorizzare (pasteurize) derivation Pasteur**

In these cases *derivation* fits well because it is a morphological relation which links the proper name with its derivatives and viceversa. As in EWN, it is used to encode derivation links when no other semantic relation is available.

In the case of metaphor, instead, there is a substitution on the basis of similarity, and, like similitude, sentences like ‘she is a true Venus’ are not reversible.

### 3. Lexicographic Experiences

#### 3.1 WordNet 2.0

The presence of a small subset of pn (Adonis, Waterloo, Cinderella, Casanova, Peter Pan, Eden, Cashmere, Champagne, etc.) has been compared in WN 1.5 and in WN 2.0; it has been verified that the same concepts are also present in English (as in Italian) and that they are far more numerous in WN 2.0 than in WN 1.5. We can say, therefore, that they are taken in greater consideration than before; moreover, each pn of this subset is present with two (or more) senses, showing the same extension of meaning as in Italian from the literal to the metaphoric or metonymical sense, even if not codified by means of a relation. E.g.:

The noun "champagne" has 2 senses in WordNet.:

1. champagne, bubbly -- (a white sparkling wine either produced in Champagne or resembling that produced here)
2. Champagne, Champagne-Ardenne -- (a region of northeastern France)

The noun "marathon" has 3 senses in WordNet.

1. marathon, endurance contest -- (any long and arduous undertaking)



A	B	C	D	E	F	G	H
Name	Frequency	Proper use	=C/B %	Extended	=E/B %	Other	=G/B %
Quirinale	487	141	28,95%	335	68,79%	11	2,26%
Palazzo Madama	211	30	14,22%	181	85,78%	0	0,00%
Picasso	154	129	83,77%	11	7,14%	14	9,09%
Caravaggio	75	16	21,33%	54	72,00%	5	6,67%
Babele	70	40	57,14%	18	25,71%	12	17,14%
Casanova	65	12	18,46%	13	20,00%	40	61,54%
Chianti	61	52	85,25%	9	14,75%	0	0,00%
Rambo	59	10	16,95%	33	55,93%	16	27,12%
Doppler	56	0	0,00%	56	100,00%	0	0,00%
Caporetto	51	30	58,82%	21	41,18%	0	0,00%
Dedalo	32	2	6,25%	16	50,00%	14	43,75%
Alzheimer	30	19	63,33%	9	30,00%	2	6,67%
Mecenate	27	2	7,41%	20	74,07%	5	18,52%
Pulitzer	20	9	45,00%	8	40,00%	3	15,00%
Stradivari	13	8	61,54%	3	23,08%	2	15,38%
Telemark	3	0	0,00%	3	100,00%	0	0,00%
<b>Total</b>	<b>1414</b>	<b>500</b>	<b>35,36%</b>	<b>790</b>	<b>55,87%</b>	<b>124</b>	<b>8,77%</b>

Table I – Statistics from PAROLE corpus

A	B	C	D	E	F	G	H
Name	Frequency	Proper use	=C/B %	Extended	=E/B %	Other	=G/B %
Quirinale	1023	372	36,36%	601	58,75%	50	4,89%
Palazzo Madama	428	23	5,37%	379	88,55%	26	6,07%
Picasso	300	193	64,33%	60	20,00%	47	15,67%
Caravaggio	185	129	69,73%	31	16,76%	25	13,51%
Babele	138	22	15,94%	97	70,29%	19	13,77%
Casanova	94	25	26,60%	27	28,72%	42	44,68%
Chianti	150	40	26,67%	108	72,00%	2	1,33%
Rambo	68	20	29,41%	39	57,35%	9	13,24%
Doppler	10	0	0,00%	10	100,00%	0	0,00%
Caporetto	45	15	33,33%	30	66,67%	0	0,00%
Dedalo	58	13	22,41%	39	67,24%	6	10,34%
Alzheimer	126	0	0,00%	117	92,86%	9	7,14%
Mecenate	77	8	10,39%	57	74,03%	12	15,58%
Pulitzer	23	2	8,70%	21	91,30%	0	0,00%
Stradivari	14	1	7,14%	11	78,57%	2	14,29%
Telemark	10	0	0,00%	9	90,00%	1	10,00%
<b>Total</b>	<b>2749</b>	<b>863</b>	<b>31,39%</b>	<b>1636</b>	<b>59,51%</b>	<b>250</b>	<b>9,09%</b>

Table II – Statistics from CLIC corpus

#### 4. A New Semantic Relation

This new research on the corpus, confirmed the growing importance of the sense extensions in the common sense expressions of every day life. With the purpose of taking in due consideration this fact we codified a new semantic relation in IWN (Marinelli R., 2004). When there is a regular shifting from a class to another belonging class (either in the case of metonym or in the case of metaphor), also for **pn** we indicate the regular shifting using the code: '*has extension*' and its reversed '*is extension of*' e.g.:

Quirinale1 *belongs to class* palazzo (palace)  
**Quirinale2** *belongs to class* carica (office)  
**Quirinale1** *has extension* **Quirinale2**  
**Quirinale2** *is extension of* **Quirinale1**

Making explicit these sense extensions for pn was a useful improvement for the IWN database. Our proposal is to study more in detail the various subjects described, continuing our research in the cognitive linguistics field exploiting this relation also for the other grammatical categories. Many more connections will be created that may teach us about mechanisms of metaphor production and comprehension (Fellbaum 2004), considering that 'the structures underlying the distinct meanings of the words are the heart of the cognitive linguistics enterprise' (Kilgariff, 1997). The results of our research in this field will be fostered by the analysis and comparison with our linguistic resources.

## References

- Copestake A., Briscoe E.J., Semi-productive polysemy and sense extension. *Journal of semantics*, 12, 15-67, 1995.
- Fellbaum, C. *Metaphors in the (mental) lexicon*, Proceedings of GWC 2004 Second Global WordNet Conference, Brno, 2004.
- Kilgariff A., Thesis: "Polysemy" University of Sussex, 1992.
- Kilgariff A., I don't believe in word senses, in *Computers and the Humanities* 31 (2), pp 91-113-University of Brighton ITRI-97-12 July, 1997.
- Marinelli R., Biagini L., Bindi R., Goggi S., Monachini M., Orsolini P., Picchi E., Rossi S., Calzolari N., Zampolli A.,: "The Italian Parole Corpus: an Overview". In: *Linguistica Computazionale*, Giardini Editori, Pisa, pp. 401-422, 2003.
- Marinelli R., Roventini A., Proper Names in a Semantic Database, Proceedings of LREC 2002 Third International Conference on Language Resources and Evaluation, Las Palmas de Gran Canaria, 2002.
- Marinelli R. 2004. Proper Names and Polysemy: from a Lexicographic Experience in: LREC 2004: Fourth International Conference on Language Resources and Evaluation, Lisbon, Portugal, 26-27-28 May 2004, Proceedings, Volume I, Lisbon, The European Languages Resources Association (ELRA). 157,160.
- Nunberg G.D., Transfers of meaning, *Journal of Semantics*, 12, 1995.
- Papafragou A., Figurative language and the semantics-pragmatics distinction, *Language and Literature*, 5, 179-93, 1996.
- Papafragou A., Metonymy and relevance, *UCL Working Papers in Linguistics*, 7; 141-175, 1995.
- Peters W., Extraction of implicit Knowledge from WordNet, Proceedings of Ontolex2002 Workshop on Ontologies and Lexical Knowledge Bases, preceding LREC2002, Las Palmas, 2002.
- Pustejovsky J. "The generative lexicon", MIT Press, Cambridge Mass. 1995.
- Pustejovsky J., Generativity and Explanation in Semantics. In: Bouillon P. and Busa F. *The Language of Word Meaning*, Cambridge University Press, Cambridge, MA, 2001.
- Roventini A., Alonge A., Bertagna F., Calzolari N., Marinelli R., Magnini B., Speranza M. (2002). "ItalWordNet: a Large Semantic Database for the Automatic Treatment of the Italian Language" in: Proceedings of the First Global WordNet Conference, Central Institute of Indian Languages, Mysore, India, pp.1-11.
- Vossen P., Condensed Meaning in EuroWordNet. In: Bouillon P. and Busa F. *The Language of Word Meaning*, Cambridge University Press, Cambridge, MA, 2001.