

# An XML Annotation Scheme for Space in an Italian Corpus

Linda Meini, Giovanna Marotta, Alessandro Lenci, Margherita Donati

Department of Linguistics, University of Pisa

Via Santa Maria 36, 56126 Pisa (Italy)

linda.meini@for.unipi.it, gmarotta@ling.unipi.it, alessandro.lenci@ling.unipi.it, margherita-donati@libero.it

## Abstract

The new resource we present consists of a corpus of oral spatial descriptions performed by congenital blind and sighted Italian subjects. The collection of the data is part of a wider project on semantic representations in the language of the blind, carried out at the Department of Linguistics, University of Pisa. The long term goal of the project is to use the evidence collected on congenital blind subjects to get at a better understanding of the relationship between linguistic and perceptive information. The corpus is currently being enhanced with different layers of annotation, focusing on spatial information. The annotation allows us to highlight the effect of the specific lexical and grammatical features of Italian on the encoding of space (e.g. with respect to the way spatial relations are encoded in motion verbs). Our resource is not only one of the few annotated corpora of spoken Italian, but it is also the first one that focuses on spatial categories.

## 1. Introduction

Space has a fundamental role in human thinking and reasoning. Like time, it is clearly a core domain of human cognition, hence it represents an ideal testing ground for an in-depth analysis of the dynamic interplay between language and non-linguistic cognition. The relationship between cognitive representations and external reality is not trivial and it gets even more complex when spatial linguistic categories are factored in (for an introduction, see Marotta, 2010).

In linguistics, a longstanding debate exists between at least two main alternative models of the relationship between language and concepts.

According to a ‘*nativistic*’ approach, the structures of spatial language are determined by our pre-linguistic categorization of space. The idea is that there is a restricted list of primitive, universal and innate notions, shared by all human beings; these notions include mostly topological relations (i.e. containment, support, contact, and proximity), and are mapped more or less directly into adpositions (Piaget and Inhelder, 1948; Miller and Johnson-Laird, 1976; Slobin, 1985; Talmy, 2000; Jackendoff, 2002).

However, scholars who analyzed spatial reference in a cross-cultural and cross-linguistic perspective have shown that both the kind of spatial relations encoded in language and the grammatical classes encoding spatial relations can vary dramatically from what we find in Western languages. These studies raised the question of how such linguistic variety can be found and accounted for if all human beings start with the same set of primitives. A new ‘*relativistic*’ approach has been developed. According to various scholars spatial language is conditioned in several ways and to several degrees by cultural conventions, and reflects representations created by exposure to spatial words relating to one’s native language. In brief, the structure and the lexicon of spatial language constrain the shape and the categories of “spatial thought” (Levinson 2003; Levinson and Wilkins, 2006; Landau *et al.* 2010).

The relationship between linguistic and non-linguistic categories grows in complexity when the reality ‘out-there’ is taken into account. Cognitive approaches assume that meanings coded by human language reflect reality as it is experienced by human beings. That is, our representation of reality is mediated by both the sensorimotor abilities of our bodies and the mental processes (basically automatic and unconscious) that organize perceptive stimuli. Therefore, some scholars (e.g. Talmy, 1983; Herskovits, 1986; Vandeloise, 1991) claimed that entities involved in spatial descriptions are not real objects, but rather geometrical abstractions of real entities that speakers conceptualize as points, lines, surfaces, or volumes. In addition, according to Vandeloise (1991) these geometrical abstractions are associated with prototypical functions that reflect how objects act in the world out there, and how we interact with them: for instance, a bowl is conceptualized as a volume with a containment function, which is coded in English by the preposition *in* ‘in’. As for dynamic descriptions, Talmy (1983) has proposed various primitive templates or ‘schemas’ for representing motion. For instance, a moving object may be described as a geometric point moving along a path - that is a line - and/or towards another object conceptualized as a point: e.g. *the ball rolled along the ledge (toward the lamp)*.

The project currently being developed at the Department of Linguistics, University of Pisa<sup>1</sup>, aims at carrying out a comparative analysis of semantic representations in congenital blind subjects with respect to those of sighted subjects. The language and the conceptual structures of blind subjects have an inherently different experiential base, which is not grounded on the visual modality. Therefore, the semantic analysis of the language of congenital blinds can provide new insights on the important relationship between conceptual structures and sensory-motor information, and more in general on the relationship between language and experience.

---

<sup>1</sup> The project we are referring to is the PRIN project 2008-2010, number 2008CM9MY3.

## 2. Linguistic Specifications

The relationship between physical and spatial properties of the world ‘out there’, human cognition, and language is very complex. Landau *et al.* (2010) have recently emphasized that language is inherently selective, encoding certain distinctions and not others; this property enables language both to modulate attention and to serve as a mental pointer, indicating which of many possible representations we have in mind.

We would like to point out that two types of selectivity can be recognized: one depending on the speaker and the other depending on the structure of language. The first type is related to what the speaker wants to communicate and how (s)he conceptualizes a given scene in a given moment. In his studies on spatial language, Talmy (1983: 225) introduces the notion of ‘schematization’, i.e. “a process that involves the systematic selection of certain aspects of a referent scene to represent the whole, while disregarding the remaining aspects”. Within a functionalist approach, some years later, Tyler and Evans (2003: 53) proposed the notion of ‘vantage point’, that “suggests that how a particular spatial scene is viewed will in large part determine the functional nature of a particular spatial scene”. In other words, spatial relations between entities are not fixed once and for all, rather they largely depend on the speaker’s perspective (Vandeloise, 1991: 23). Therefore, the linguistic description of a spatial scene is shaped by the specific speaker’s point of view and his/her communicative purposes.

The second type of selectivity is related to the structure of language. Spatial relations are usually encoded by some grammatically defined classes:

- verbs of position and motion: e.g. Eng. *lie, sit, stand, roll, arrive, reach, go*;
- adpositions and particles or adverbs: e.g. Eng. *in, on, at, across, up*.

Languages typically lexicalize in each grammatical class specific semantic content, which varies from language to language.

For instance, languages differ with respect to which semantic components are lexicalized in the class of motion verbs (Talmy, 1985; Slobin, 2004). In English verbs typically lexicalize the conceptual components of Motion and Manner (e.g. *to roll, slid, walk, run*), while Path is encoded out of the verb by prepositions or particles (e.g. *to rolled off, walk into, go up*). By contrast, Italian verbs mostly lexicalize the conceptual components of Motion and Path (e.g. *uscire* ‘to go out’, *entrare* ‘to go in’, *salire* ‘to go up/get on’), while Manner is optionally encoded out of the verb by adverbials or gerundive type constituents (e.g. *Sei salito a piedi?* lit. ‘did you go up (on foot)’, i.e. ‘did you walk up?’)<sup>2</sup>. There are other possibilities. Languages like Atsugewi (a Californian Indian language, now extinct) have a whole series of verb roots that lexicalize Motion and various kind of objects or

<sup>2</sup> Languages that behave like English (e.g. German) are classified as *Satellite-Framed*, whereas languages like Italian (e.g. other Romance languages) are called *Verb-Framed* (Talmy, 1991).

materials as moving and located: e.g. *-lup-* ‘for a small shiny spherical object to move/be-located’, *-qput-* ‘for loose dry dirt to move/be-located’ (examples from Talmy, 1985). Languages can also differ with respect to which spatial relations are encoded by prepositions. For instance, Italian *su* can encode all the relations that in English are expressed by *on (upon/onto), over, above, on top of, up*. In fact, it is well known that spatial prepositions are ambiguous and highly context dependent (Vandeloise, 1991; Tyler and Evans, 2003; Meini, 2009). Another way to encode spatial relations is found, again, in Atsugewi. It has a set of verbal suffixes that encodes the Path and the type of objects or materials where the motion is directed to: e.g. *-içt* ‘into a liquid’, *-mic* ‘down into (/onto) the ground’ (examples from Talmy, 1983).

The previous examples showed clearly that language’s spatial system imposes a fixed form of structure on virtually every spatial scene. In other words, speakers cannot describe a spatial scene in just any way they might wish, rather they must choose among the word classes available in the organization of the lexicon of their specific language.

## 3. Data Collection

Within our project, 22 congenital blind subjects were selected, 10 females and 12 males, ranging from 21 to 72 years old (female average age: 47; male average age: 45). Of these 22 subjects 12 are from Tuscany, 5 from Liguria and 5 from Sardinia. On the basis of their personal data (such as age, gender, city of residence), as well as socio-linguistics parameters<sup>3</sup>, 22 corresponding sighted subjects have been selected, with characteristics similar to those of the blind people.

Three spatial tasks were submitted to all the informants in a randomized order. The tasks have been designed with the following aims:

- eliciting spatial descriptions comparable among subjects;
- eliciting static as well as dynamic descriptions;
- displaying various situations, that might elicit different perspectives or Frames of Reference<sup>4</sup>;
- selecting places according to different degrees of familiarity.

Therefore, we designed the following (semi-spontaneous) tasks:

*Task a.: bedroom description* - The subject is asked to describe his/her own bedroom as thoroughly as possible. After the subject has completed the description, the interviewer asks some common questions about the

<sup>3</sup> We took into account not only the educational attainment but also speakers’ culture in general, assessed by the number of books they read, the movies they watch and their hobbies. The kind of place they live in (city or small town or even countryside), their ‘social’ life (whether they have friends and go out with them) and the degree of their autonomy (e.g. whether they need to be accompanied or supervised by someone else) have been regarded as well.

<sup>4</sup> For the notion of Frame of Reference, see § 4.2.2.

bedroom, e.g. *Is there any picture on the wall? Where are they? Where is the bed with respect to the door?* The answers are meant to provide us with both quantitatively and qualitatively similar data across different subjects. The bedroom is a familiar place, which is meant to elicit a (mostly) static description.

*Task b.: urban itinerary* - With the modality of role play, the interviewer pretends to be a tourist who meets the subject at a specific point A of his/her city and asks him/her how to reach a specific point B. The subject is also asked to take a route such that the tourist would be able to see as many touristic places as possible. After the subject has completed the description, the interviewer asks some common questions about the chosen route, e.g. about the route in Pisa: *Where is piazza dei Cavalieri with respect to piazza Santa Caterina? How far is piazza dei Miracoli from piazza Santa Caterina?* We selected two routes, one 'easier' (allegedly best known) than the other. In either cases the chosen itineraries were supposed to be known by the subjects. However, some blind subjects did not know either routes, because of their lack of movement autonomy. Therefore, in some cases the subject himself/herself proposed a short route he/she covers daily. The task proposed is meant to elicit a (mostly) dynamic description.

*Task c.: bird-eye city description* - The subject is asked to describe his/her city from the highest point of view possible (e.g. a bell-tower) to a tourist who has never visited it. After the subject has completed the description, the interviewer asks some common questions about the city, e.g. about Pisa: *Where is piazza del Duomo with respect to the rail station? How far are they? Is it big? What shape do you think it is?* One's own city is supposed to be a known place, but the point of view is completely unusual for both blind and sighted subjects. The task proposed should elicit a (mostly) static description and provide data to contribute to the current debate on whether blind individuals show some 'preference' for a specific spatial perspective (e.g. route over survey: see Taylor and Tversky 1992, 1996) or reference frame (Noordzij *et al.*, 2006) opposite to that chosen by sighted people.

All tasks, submitted and performed orally, were recorded. The audio files were then transcribed, using the Dragon speech recognition software with a re-speaking technique, and then manually checked. Then the transcriber, while listening again to the audio files, corrected manually all the inconsistencies and misspellings in the transcripts. Finally, the transcripts have been checked by the interviewer who actually submitted the tasks.

The transcript format used is CHAT, by the CHILDES project (MacWhinney, 2000). Although we are aware of other formats (e.g., annotation graphs; Bird and Liberman, 2001), developed especially in the area of multimodal annotations, we chose CHAT to transcribe the audio files as it is the current standard transcription system used in psycholinguistic analyses.

Figure 1 represents an example of a transcript according to the CHAT coding scheme: it is an excerpt from task a. (It. *descrizione stanza*, 'room description') performed by a blind (It. *non vedente*) male subject from Lucca (LU).

```
Non_Vedente_A.R._LU

i@Begin
@Languages: ita
@Participants: SUB Subject13nonvedente, INV Investigator
@ID: ita|descrizione stanza|SUB|36;9.25|male|||Subject13|18|
@ID: ita|descrizione stanza|INV||female|||Investigator||
@Date: 06-APR-2011
@Transcriber:Giulia
*INV: Ok, quindi descrizione della stanza propria.
Descrivi nella maniera più dettagliata possibile la tua camera da letto.
*SUB: Quindi compresi gli oggetti, la posizione degli oggetti?
*INV: Sì, tutto quello che vuoi.
*SUB: Ok, no vabbe' chiedo perché+//.
Allora intanto diciamo che la forma della stanza (...) è (eee), almeno dal punto di vista percettivo, rettangolare.
(...) è una: stanza se- (mmm), diciamo così due caratteristiche, abbastanza grande per cui viene divisa in: [x2] due.
```

Figure 1: Excerpt of a transcript in CHAT

#### 4. Description of our Annotation Scheme

Bearing in mind the goals of our project (§ 1), we designed an XML annotation scheme able to capture both the specific structure of the language used by the informants (viz. Italian) and the underlying conceptual components or strategies that yield a specific spatial description (§ 3). So our approach is basically empirical, driven by actual instances of language use found in our corpus.

For the moment we decided not to adopt existing annotation languages, such as ISO-Space for instance (Pustejovsky *et al.* 2011). ISO-Space is an annotation specification, designed for capturing spatial and spatio-temporal information in natural language text. It aims at providing an inventory of how spatial information is presented in natural language such that it can be integrated by complement information coming from other modalities (e.g. GPS). ISO-Space is supposed to serve different purposes, such as, for instance: determining object location given a verbal description, constructing a route given a route description, integrating spatial descriptions with information from other media, reconstructing spatial information associated with a sequence of events, etc. (Pustejovsky *et al.* 2011: 1). We would like to briefly present two characteristics of ISO-Space, to highlight the differences between our annotation scheme and that of ISO-Space. To serve the above-mentioned purposes, ISO-Space distinguishes two major types of elements: entities (that include location and spatial entities, as well as both dynamic motion and static arrays) and spatial relations (that specify what kind of relation holds between the entities involved). As for locations, the annotation scheme of ISO-Space provides a number of elements and attributes that can be easily integrated with information deriving from other resources, such as physical feature databases and gazetteers (Pustejovsky *et al.* 2011: 3). As

for motion verbs, ISO-Space uses a classification by Muller (1998), that distinguish verbs according to their semantic meaning: *move*, *move\_external* (if the motion takes place outside), *move\_internal* (if the motion takes place inside), *leave*, *detach*, *deviate*, etc.

By contrast, in designing our annotation scheme, focused on the analysis of linguistic spatial categories of Italian, we adopted many distinctions commonly drawn in linguistic studies on the encoding of space, that are not included in ISO-Space (or are used differently: see, e.g., “qualitative spatial links”). Therefore, our annotation scheme is designed to analyze how language encodes space, whereas ISO-Space is designed to explicit how the space ‘out-there’ is captured by language.

However, with our markup language we would like to contribute to the current development of other annotation languages, such as ISO-Space, from a ‘more linguistic’ point of view (cf. Mani and Pustejovsky, 2012).

The XML markup language we designed encodes both morpho-syntactic and semantic categories. So far 88 recordings, corresponding to task a. and b., have been annotated by three annotators.

In this section we describe the main characteristics of our markup language, using example annotations from our corpus.

Our markup language provides two major tags:

- `<motion_event>`;
- `<localization_event>`.

In other words, there is a major distinction between dynamic and static descriptions.

#### 4.1 Motion event

It is every situation involving either movement or displacement<sup>5</sup>.

In linguistics, dating back to Tesnière (1959), scholars usually distinguish between ‘movement’ and ‘displacement’. The first term refers to the type of motion encoded by the verb (e.g. *to roll*, *slid*, *walk*, *run*); the second to the displacement, i.e. the complete shift of an entity through space. Our definition of ‘motion event’ embraces both distinctions. We included metaphorical motion as well: it is marked with the attribute “fictive”<sup>6</sup>:

```
<motion_event type="fictive">
  Via Duomo ci porterebbe in piazza del Giglio7
</motion_event>
```

Every *motion\_event* element usually contains other two tags:

- `<motion_verb>`;
- `<spatial_role>`.

<sup>5</sup> Please notice that our definition of ‘motion event’ differs from that by Talmy (1985), which includes both motion and stationary location.

<sup>6</sup> “Linguistic instances that depict motion with no physical occurrence” (Talmy 2000, I: 99).

<sup>7</sup> ‘Via Duomo would take us in piazza del Giglio’.

#### 4.1.1. Motion Verbs

By means of different attributes, we distinguish between different types of verbs. These distinctions rely mostly upon the conceptual components identified by Talmy (1985: 61): Motion: “refers to the presence *per se* in the event of motion”; Path: “is the course followed ... by the Figure object with respect to the Ground object”<sup>8</sup>; Manner: refers to the type of motion. Therefore, the attributes of the element *motion\_verb* are the following:

- **dislocation**: the verb lexicalizes only the Motion:

```
<motion_event>
  <motion_verb type="disl">Si va </motion_verb>
  in piazza San Martino9
</motion_event>
```

- **path**: the verb lexicalizes the Motion and the Path:

```
<motion_event>
  <motion_verb type="path">Attraversi </motion_verb>
  il ponte10
</motion_event>
```

- **manner**: the verb lexicalizes the Motion and its Manner:

```
<motion_event>
  Io continuo a
  <motion_verb type="manner">camminare </motion_verb>
  sul marciapiede11
</motion_event>
```

- **conveyance**: the verb conflates the self-movement of the “carrier” entity and the “caused-movement” of the entity “carried” (e.g. *to bring*, *take*, *lead*)<sup>12</sup>:

```
<motion_event type="fictive">
  Via Duomo ci
  <motion_verb type="conv">porterebbe </motion_verb>
  in piazza del Giglio13
</motion_event>
```

The last two attributes rely not on the semantics of the verbs, but on their syntactic encoding:

- **construction**: the verb itself does not encode the notion of Motion, which is conveyed by the pair verb-direct object instead:

```
<motion_event>
  <motion_verb constr="y">fare </motion_verb>
  quindi tutta la via Grande14
</motion_event>
```

<sup>8</sup> “The Figure is a *moving* or conceptually *movable* object whose path or site is at issue; the Ground is a reference-frame, or a reference-point stationary within a reference-frame, with respect to which the Figure’s path or site is characterized” (Talmy, 1983: 232).

<sup>9</sup> ‘We go in piazza San Martino’.

<sup>10</sup> ‘You cross the bridge’.

<sup>11</sup> ‘I keep walking on the sidewalk’.

<sup>12</sup> Many scholars limit their research to intransitive verbs that encode human prototypical motion.

<sup>13</sup> For the translation, see footnote n. 7.

<sup>14</sup> Lit. ‘(we have) to do all via Grande’, i.e. ‘we have to pass through via Grande’. Besides path constructions, we found manner constructions as well: e.g. *farmi una passeggiata*, lit. ‘I do (i.e. take) a walk’.

- **phrasal verb**: it is a phrase constituted by a verbal head and a complement represented by a ‘particle’ (originally an adverb); its syntactic cohesion is so tight that it is not possible to replace the whole phrase with only one of its parts<sup>15</sup>:

```
<motion_event type="fictive">
  dal cimitero
  <motion_verb phv="y">andava su</motion_verb>16
</motion_event>
```

#### 4.1.2. Spatial roles

In the literature concerning motion events, there are distinctions relative to the entities that are used as the reference-frame of the events. These distinctions have been introduced by Fillmore (1971, now 1997: 40): when talking about ‘locomotion’ (i.e. an object change of location in time), he formulates the “case-like” notions of *source*, *goal*, *path* and *location*. They do not represent conceptual elements, like Talmy’s notions (§ 4.1.1), but thematic roles: therefore they refer to the semantic function of a noun phrase with respect to its verb.

Four main distinctions pertain to spatial roles, that are marked also by the attribute indicating the part of speech (pos), such as prepositional phrase (pp), noun phrase (np), adverb (adv), etc.:

- **source**: is the place whence an entity departs:

```
<motion_event type="fictive">
  <spatial_role role="source" pos="pp">dal cimitero </spatial_role>
  <motion_verb phv="y">andava su</motion_verb>17
</motion_event>
```

- **goal**: is the destination reached by an entity:

```
<motion_event type="fictive">
  Via Duomo ci
  <motion_verb type="conv">porterebbe </motion_verb>
  <spatial_role role="goal" pos="pp">in piazza del
  Giglio</spatial_role>18
</motion_event>
```

- **path**: is the course followed by an entity during its motion (cf. Talmy’s definition, § 4.1.1):

```
<motion_event>
  <motion_verb type="path">Attraversi </motion_verb>
  <spatial_role role="path" pos="np">il ponte</spatial_role>19
</motion_event>
```

- **manner**: it is the manner of motion<sup>20</sup>:

```
<motion_event>
  <motion_verb type="path">Sono partito </motion_verb>
  <spatial_role role="manner" pos="adv">di scatto </spatial_role>21
</motion_event>
```

<sup>15</sup> The definition is adapted from Simone (1996: 49), where it is applied to the Italian ‘verbi sintagmatici’ (‘syntagmatic verbs’).

<sup>16</sup> ‘From the graveyard it [*scil.* the road] went uphill’.

<sup>17</sup> For the translation, see footnote n. 16.

<sup>18</sup> For the translation, see footnote n. 7.

<sup>19</sup> For the translation, see footnote n. 10.

<sup>20</sup> We introduced this new spatial role because in the Verb-framed languages (such as Italian), the Manner is usually encoded - when encoded at all - by an adjunct (§ 2).

<sup>21</sup> Lit. ‘I left suddenly’, i.e. ‘I bolted’.

## 4.2 Localization event

It is every situation involving a stationary location of an entity (Figure) with respect to other entities (Ground).<sup>22</sup>

Every *localization\_event* element usually contains other two elements:

- <localization\_verb>;
- <spatial\_relation>.

### 4.2.1. Localization Verbs

Many languages have a series of verbs that describe the static position of a Figure with respect to a Ground entity (cf. Ameka and Levinson, 2007). See, for instance, the following postural verbs: En. *to lie, sit, stand, hang*; Ger. *liegen* ‘to lie’, *sitzen* ‘to sit’, *stehen* ‘to stand’, and their respective dynamic counterpart *legen* ‘to lay’, *setzen* ‘to sit’, *stellen* ‘to stand’ (Rüsch, 2010).

In Italian there are very few verbs of position and they are rarely used: e.g. *giacere* ‘to lie’ (which pertains to the literary register), *sedere* ‘to sit’. In every day communications, periphrastic constructions are preferred: they are formed by the verbs *stare* ‘to stay’ or *essere* ‘to be’, and a former past participle or an adverbial: *stare sdraiato* ‘to lie’, *stare seduto* ‘to sit’, *stare in piedi* ‘to stand’, *essere/stare appeso* ‘to hang’. The spatial information conveyed by these verbs and constructions in Italian is optional; moreover it prototypically pertains to human beings.

Since the semantics of Italian localization verbs does not entail many distinctions, at the moment our markup language does not provide further attributes for localization verbs<sup>23</sup>. In fact, in our corpus we found mostly the verb *essere/esserci* ‘to be/there be’ or verbs like *avere* ‘to have’, and *trovare/trovarsi* ‘to find/be’:

```
<localization_event>
  Su questo mobiletto
  <localization_verb>c’è </localization_verb>
  un cestino
</localization_event>24
```

### 4.2.2. Spatial Relations

In the Western tradition of spatial studies, a lot of attention has been paid to (static) spatial relations, at least since the pioneer work by Piaget and Inhelder (1948) on the development of spatial representations in human beings. They showed that the first spatial concepts acquired are the topological notions of proximity, order, closure, and continuity; only much later, children understand the Euclidean notions of metric distance and

<sup>22</sup> Linguistic description of space, both static and dynamic, is highly relational (Meini, 2009; Mani and Pustejovsky, 2012): usually we locate an entity x (Figure) by reference to the location of entity y (Ground). Similarly, we describe an entity w (Figure) as moving in relation to an entity z (Ground); a dynamic description is not relational when it involves only the ‘movement’ of the Figure (§ 4.1).

<sup>23</sup> However, distinctions concerning localization verbs (even postural verbs for the annotation of languages different from Italian) could be easily added into our markup language.

<sup>24</sup> ‘On this little table there is a bin’.

angle, and, at last, they are able to grasp geometrical projective relations.

Irrespective to the different and opposing theoretical paradigms developed on the basis of Piaget and Inhelder's findings (§ 1), the distinction between topological and projective relation is still maintained in linguistic researches<sup>25</sup>. Therefore, in our markup language, we distinguish these two main types of spatial relations<sup>26</sup>. Each entity involved in both topological and projective relations are marked by the attribute indicating the part of speech (§ 4.1.1).

Following studies on spatial relations (Becker, 1997; Meini, 2009) we distinguish six main kinds of TOPOLOGICAL relations<sup>27</sup>:

- **at place**: the Figure is localized with respect to the Ground entity without any further spatial information:

```
<spatial_relation type="top" subtype="at place" pos="pp">
  alla finestra28
</spatial_relation>
```

- **inner**: the Figure is localized with respect to the inner subspace of the Ground entity:

```
<spatial_relation type="top" subtype="inner" pos="pp">
  nella mia stanza da letto29
</spatial_relation>
```

- **neighbouring**: the Figure is localized with respect to the subspace surrounding the Ground entity:

```
<spatial_relation type="top" subtype="neigh" pos="pp">
  vicino al letto30
</spatial_relation>
```

- **boundary**: the Figure is localized with respect to the boundary of the Ground entity:

```
<spatial_relation type="top" subtype="boundary" pos="pp">
  Su questo mobiletto31
</spatial_relation>
```

- **exterior**: the Figure is localized with respect to the exterior subspace of the Ground entity:

```
<spatial_relation type="top" subtype="ext" pos="pp">
  esternamente alla portafinestra32
</spatial_relation>
```

- **between**: the Figure is localized with respect to a 'complex' Ground composed of disjunct entities:

```
<spatial_relation type="top" subtype="betw" pos="pp">
  tra il letto e la finestra
</spatial_relation>33
```

PROJECTIVE relations suppose the notion of direction and a system of axis. Therefore, we distinguish three main kinds of projective relations, according to the axis involved:

- **lateral**: <spatial\_relation type="proj" axis="lat" pos="pp">  
sulla destra<sup>34</sup>  
</spatial\_relation>
- **sagittal**: <spatial\_relation type="proj" axis="sag" pos="pp">  
di fronte al campo da tennis<sup>35</sup>  
</spatial\_relation>
- **vertical**: <spatial\_relation type="proj" axis="ver" pos="pp">  
sopra il letto<sup>36</sup>  
</spatial\_relation>

For every projective relation we also indicate the frame of reference (henceforth FoR)<sup>37</sup>.

The notion of FoR has been introduced in linguistic analyses by Levinson (2003: 24), who provides the following definition, quoting from Irvin Rock: "a unit or organization of units that collectively serve to identify a coordinate system with respect to which certain properties of objects, including the phenomenal self, are gauged"<sup>38</sup>. Three main FoR have emerged from Levinson (2003: 38 ff.) analysis (the corresponding attribute of our annotation is in brackets):

- **Intrinsic (FoR="intr")**: the coordinates are determined by the so-called 'inherent features' (i.e. sidedness or facets) of the Ground entity;
- **Relative (FoR="rel")**: directions are assigned to Figure and Ground by the coordinates fixed on a distinct 'viewpoint';
- **Absolute (FoR="absol")**: this FoR is based on fixed bearings, such as cardinal points.

Sometimes linguistic data are not enough to tell an intrinsic from a relative FoR, since in many languages there are not linguistic items specific for the two domains. In these cases, the annotator cannot only rely on the linguistic text, but has to supplement it with extra-textual information.

Since the kind/s of FoR usually employed in a language is/are culturally determined (Levinson, 2003), the annotation will allow us to analyze which FoR is mostly used (or mostly avoided) by Italian speakers. We will also evaluate whether the choice of a specific FoR is influenced by the kind of spatial description, e.g. a small

<sup>25</sup> For more recent studies on the development of prelinguistic spatial concepts (and its relationship with the acquisition of spatial language), see e.g. Mandler (2004).

<sup>26</sup> For a discussion of how Italian, English and Spanish encode both topological and projective relations by means of prepositions or adverbials, see Meini (2009).

<sup>27</sup> For reasons of space, in this section we quote only the annotation relative to the spatial relation and not that of the whole localization event. In the translation, the full sentence is provided.

<sup>28</sup> '(There are curtains) at the window'.

<sup>29</sup> 'In my bedroom (there is a door)'.

<sup>30</sup> '(it [*scil.* the armchair] is) near the bed'.

<sup>31</sup> For the translation, see footnote n. 24.

<sup>32</sup> 'Outside the French doors (there is also a terrace)'.

<sup>33</sup> 'Between the bed and the window (there is an armchair)'.

<sup>34</sup> '(There is a television) on the right'.

<sup>35</sup> 'Opposite to the tennis court (there is shop)'.

<sup>36</sup> '(This fan is) above the bed'.

<sup>37</sup> For obvious reasons, the specification of the *axis* is not needed in case of an absolute FoR.

<sup>38</sup> It follows that FoR must not be confused with the (kind of the) origin of the coordinate system: e.g. the opposition 'egocentric' vs 'allocentric'.

room vs wide spaces (such as a city). Finally, the analysis of the FoR in our corpus will reveal whether there are any differences between blind and sighted individuals, hence whether the different experiential base influences somehow the choice of the FoR.

As for our annotation, we drew a sketch of the speakers' bedrooms (task a., Figure 2), and we checked the route followed by the informants with Google Map or Google Earth (task b., Figure 3); we suppose that Google Earth will be an useful tool when annotating the descriptions of task c. as well.

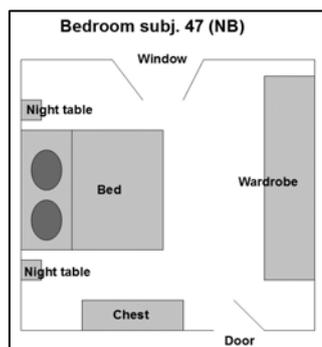


Figure 2: Sketch of subj. 47's (Not Blind) bedroom

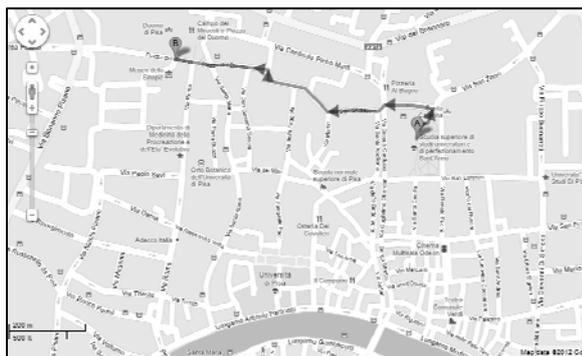


Figure 3: Route followed by subj. 23 (Google Map - Pisa)

### 4.3 Spatial Attributes

Finally, the markup language provides the element *attribute*, which concerns mostly spatial attributes, such as: shape (e.g. *ottagonale* 'octagonal'), size (e.g. *piccolo* 'small'), material (e.g. *di legno* 'wooden'). Moreover, we included the attribute *metric distance*, to verify whether the speakers' more or less accuracy in distance comparison (task b. and c.) could be related to the different experiential base of blind and sighted individuals (cf. also Noordzij *et al.* 2006).

As with spatial roles and spatial relations, the attribute elements are marked by the indication of the part of speech (§ 4.1.1, § 4.2.2).

## 5. Conclusion

In this paper we reported on a new annotated resource currently being developed to analyze spatial information in a corpus of spoken Italian. The resource consists of various material: spatial descriptions made by Italian

speakers (and the corresponding transcripts in CHAT format) are associated with sketches of rooms and maps of routes.

The spatial annotation scheme and markup language we designed aims at providing a comprehensive tool that allows the researcher to highlight:

- effects of the specific lexical and grammatical features of the language spoken by the informants (in this case Italian) on the encoding of space;
- differences in the encoding of space related to:
  - speakers' sociolinguistic variables, such as age, gender, dialect spoken, culture, etc. (§ 2);
  - different experiential base between blind and sighted individuals (§ 1).

In designing the markup language, we referred mainly to the linguistic literature on space and to psycholinguistic studies on the language of the blind. However, our approach is empirical: it means that among the overall amount of distinctions drawn in linguistic literature, we chose those relevant to the actual instances of language use found in our corpus.

In the very next future we are going to:

- enrich our markup language with the annotation of nouns, functioning especially as Ground (Herskovits, 1986; Vandeloise, 1991; Frank, 1997; Meini, 2009; Bateman *et al.*, 2010);
- enrich the semantic annotation of verbs, especially the motion verbs (e.g. Italian path verbs encode many different 'experiential' situations);
- carry out inter-coder agreement tests;
- automatically PoS-tag and lemmatize the corpus;
- explore possible synergies between our annotation scheme and ISO-Space;
- align speech recordings with transcripts and other coding layers using annotation graphs.

Our linguistic analyses aims at providing more evidence about spatial language use in Italian.

All the resources developed by the project will be publicly and freely available, and they should be of interest to a wide scientific community. The resources will be released with Creative Commons Attribution-Share Alike license (see <http://creativecommons.org/>).

## 6. References

- Ameka, F. and Levinson, S.C. (2007). Introduction. The Typology and Semantics of Locative Predicates: Posturals, Positionals, and Other Beasts. *Linguistics*, 45: pp. 847--871.
- Bateman, J.A., Hois, J., Ross, R., Tenbrink, T. (2010). A linguistic ontology of space for natural language processing. *Artificial Intelligence*, 174(14): pp. 1027--1071.
- Becker, A. (1997). Theoretical framework. In Becker, A. and Carroll, M. (Eds.), *The acquisition of spatial relations in a second language*. Amsterdam-Philadelphia: J. Benjamins, pp. 13--33.
- Bird, S. and Liberman, M. (2001). A Formal Framework for Linguistic Annotation. *Speech Communication* 33(1,2): pp. 23--60.

- Fillmore, C. (1971). *Santa Cruz lectures on deixis*. Bloomington: Indiana University Linguistics Club [now Fillmore, C. (1997). *Lectures on Deixis*. Stanford: CSLI Publications].
- Frank, A.U. (1997). Spatial ontology: A geographical information point of view. In: Stock, O. (Ed.), *Spatial and Temporal Reasoning*. Dordrecht: Kluwer Academic Publishers, pp. 135--153.
- Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge: CUP.
- Jackendoff, R. (2002). *Foundations of Language*. Cambridge, MA: MIT Press.
- Landau, B., Dessalegn, B. and Goldberg, A.M. (2010). Language and Space: Momentary Interactions. In Evans, V. and Chilton, P. (Eds.), *Language, Cognition and Space: The State of the Art and New Directions*. London: Equinox Publishing, pp. 51--77.
- Levinson, S.C. (2003). *Space in language and cognition: explorations in cognitive diversity*. Cambridge: CUP.
- Levinson, S. and Wilkins, G. (Eds.) (2006) *Grammars of Space*. Cambridge: CUP.
- MacWhinney, B. (2000), *The CHILDES Project: Tools for Analyzing Talk*. 3<sup>rd</sup> Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mandler, J. (2004). *The foundations of mind: Origins of conceptual thought*. Oxford: OUP.
- Mani, I. and Pustejovsky, J. (2012), *Interpreting Motion: Grounded Representations for Spatial Language*. Oxford: OUP.
- Marotta, G. (2010). Introduction. In Marotta, G., Lenci, A., Meini, L. and Rovai, F. (Eds.), *Space in Language 2009. Proceedings of the Pisa International Conference*. Pisa: ETS, pp. 11--25.
- Meini, L. (2009). *Dimensioni dello Spazio nelle Preposizioni. Un'Analisi Empirica sull'Italiano L2*. Pisa: PLUS.
- Miller, G.A. and Johnson-Laird, P.N. (1976). *Language and perception*. Cambridge: CUP.
- Muller, P. (1998). A qualitative theory of motion based on spatio-temporal primitives. In Cohn, A.G., Schubert L., and Shapiro, S.C. (Eds.), *KR'98: Principles of Knowledge Representation and Reasoning*. San Francisco, CA: Morgan Kaufmann, pp. 131--141.
- Nordzij, M. Zuidhoek, S. and Postma, A. (2006). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition*, 100, pp. 321--342.
- Piaget, J. and Inhelder, B. (1948). *La représentation de l'espace chez l'enfant*, Paris: Presses Universitaires de France.
- Pustejovsky, J., Moszkowicz, J. and Verhagen, M. (2011). ISO-Space: The Annotation of Spatial Information in Language. In Bunt, H.C. (Ed.), *Proceedings of the 6th Joint ISO-ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-6), January 11-12, 2011, Oxford, UK*. Oxford: University of Oxford, pp. 1--9.
- Rüsch, N. (2010). Placement events in German as a second language: *Legen, stellen, and setzen* and their static counterparts. In Marotta, G., Lenci, A., Meini, L. and Rovai, F. (Eds.), *Space in Language 2009. Proceedings of the Pisa International Conference*. Pisa: ETS, pp. 629--641.
- Simone, R. (1996). Esistono verbi sintagmatici in italiano?. *Cuadernos de Filología Italiana*, 3, pp. 47--61.
- Slobin, D.I. (1985). Crosslinguistic evidence for the language-making capacity. In Slobin, D.I. (Ed.), *The crosslinguistic study of language acquisition, vol. II*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1157--1256.
- Slobin, D.I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In Strömqvist, S. and Verhoeven, L. (Eds.), *Relating Events in Narrative, vol. II: Typological and Contextual Perspective*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 219--257.
- Talmy, L. (1983). How language structures space. In Pick, H.L. and Acredolo, L.P. (Eds.), *Spatial orientation: theory, research, and application*. New York-London: Plenum Press, pp. 225--282.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Shopen, T. (Ed.), *Language Typology and Syntactic Description*. Cambridge: CUP: pp. 57--149.
- Talmy, L. (1991). Path to realization: A typology of event conflation. In Sutton, L.A, Johnson, C. and Shields, R.B. (Eds.), *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, pp. 480--519.
- Talmy, L. (2000). *Toward a Cognitive Semantics, 2 Vols.*. Cambridge, MA: MIT Press.
- Taylor, H.A. and Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory & Language*, 31, pp. 261--292.
- Taylor, H.A., and Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory & Language*, 35, pp. 371--391.
- Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Tyler, A. and Evans, V. (2003). *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition*. Cambridge: CUP.
- Vandeloise, C. (1991). *Spatial Prepositions: A Case Study from French*. Chicago: The University of Chicago Press.