Towards a topography of grammatical meaning

Elisabetta Bonvino, Francesca Masini, Paola Pietrandrea

Università degli Studi Roma Tre, Dipartimento di Linguistica

Project Il significato grammaticale nelle strutture del discorso, Ms., July 2007

Abstract. One of the major contemporary challenges for Linguistics and Information and Communication Technologies is automatic information retrieval, which must rely upon a robust and accurate semantic annotation. In this article, we suggest that part of the meaning of discourse, namely, part of its grammatical meaning, is found in its topological structure. In order to correctly retrieve information from a text, one should detect its topological structure.

1 Introduction

The basic tenet of our approach is that grammatical meaning can be detected at all levels of linguistic structure. More precisely, we maintain that grammatical meaning can be expressed in the topological structure of discourse, which may be regarded as a particular type of construction. The following sections contain a brief discussion of these notions.

1.1 The topographical approach

We define grammatical meaning as the “non-denotative, general ... abstract and relational meaning” (Bybee et al. 1994 [1]: 5) that is conveyed by grammatical categories, such as tense, aspect, mood, etc. We share the view that, in order to be properly classified as “grammatical”, a form should display a reduction in autonomy, both at the paradigmatic and at the syntagmatic level (Lehmann 2005 [2]). Nevertheless, our definition of grammatical meaning entails that it can be in principle detected not only at the morphological level, but at all levels of linguistic structure, including discourse.

We define discourse as a proper structural level of analysis, endowed with its own units and structures. Building on the seminal work by Claire Blanche-Benveniste on spoken French (1979, 1990, 1997) [3-5], we claim that discourse is endowed with a topological structure.

By topological structure of discourse (henceforth TSD) we mean the form taken by the speech flow when it is represented in “grids”, i.e., not linearly, but on a bidimensional plane, as proposed by Blanche-Benveniste et al. (1979) [3] and subsequent works: the horizontal axis of this plane features the sequence of the positions that define the skeleton of the so-called “maximal construction” (i.e., the verb with its arguments and adjuncts); the vertical axis lists all the actual realizations within each position.
This representation accounts for the fact that the maximal construction is rarely uttered all at once. More frequently, it is gradually built by means of repetitions, hesitations, rewordings, additions, and other kinds of insistences on one or more of its positions. Far from being linguistic noise, these phenomena often happen to create patterns (or, as they are often referred to, “configurations”) that regularly recur in spoken and, we may add, also written texts. From a theoretical standpoint, these regularities may be accounted for by adopting a constructionist approach.

1.2 Configurations as “constructions”

In our view, the TSD can be considered as a new type of construction in the technical sense of constructionist approaches (cfr. Fillmore, Kay and O’Connor 1988 [6], Goldberg 1995 [7], 2006 [8]), i.e., an association of a form and a meaning. More precisely, we hypothesize that different TSDs may convey different grammatical meanings.

Assuming that TSDs are constructions is not a trivial operation. Indeed, this entails two main considerations. On the one hand, assuming that TSDs are constructions implies to acknowledge that we are in front of a new type of construction, since TSDs are bidimensional in nature. On the other hand, our assumption concerns one important question: how far can we extend the notion of construction? So far, most current constructionist works focused on canonical levels of analysis such as syntax (e.g. Goldberg 1995 [7]) and morphology (e.g. Booij 2005 [9]), as well as on idioms (e.g. Fillmore, Kay and O'Connor 1988 [6]). However, in the literature one can find few but incisive attempts to extend the notion of construction to other domains, such as for instance Östman (2005) [10] and Marandin (2006) [11].

Our contribution aims at making another step in this direction thanks to an innovative approach, which is obtained by the unprecedented combination of the methodological traditions developed by corpus-based analyses of discourse (grid representations), constructionist approaches and grammatical studies. In the following section we present some preliminary results.

2 Some preliminary results

2.1 Tense

As already noticed, a spoken text can be characterized by TSDs made up of lists of lexemes, repetitions of syntactic structures, chiastic syntactic structures and so on. These forms often happen to be associated to well established grammatical categories, such as, for instance, the tense category.

Let us draw our attention to the text in (1). This piece of text contains a configuration that can be easily detected if we transpose it into the grid representation in Fig. 1 (from Blanche-Benveniste 1993 [12]).
Le raisin est déversé dans un conquet de réception. Ce conquet de réception communique avec un appareil qui s'appelle fouloir égrappoir et cet appareil a pour but de fouler donc le raisin.
The raisin is poured in a reservoir. This reservoir is linked to a device called grape squeezer. This device is thus used to squeeze grapes.

The representation in Fig. 1 reveals the existence of a TSD characterized by three repetitions of the following syntactic structure: argument1 + predicate + argument2. The three lines are intersected by two left-slipping structures, i.e., two structures in which the argument2, given a first time after the predicate, is repeated a second time as argument1 before the next predicate (dans un conquet de reception - ce conquet de reception; avec un appareil - cet appareil).

According to Blanche-Benveniste (1993) [12], this TSD imposes, under certain syntactic conditions, a meaning of temporal succession to the event. This meaning is absent in the linear representation of the same event that makes use of relative clauses. The contrast between the two structures can be observed by comparing (1) to (2).

Le raisin est déversé dans un conquet de réception, qui communique avec un appareil qui s'appelle fouloir égrappoir, qui a pour but de fouler donc le raisin.
The raisin is poured in a conque of reception, which is linked to a device called grape squeezer, which has for purpose to squeeze the raisin.

Whereas (1) illustrates the temporal development of the vintage process, (2) merely provides a static representation of the instruments used for vintage. The only structural difference between (1) and (2) is the use of a TSD vs. relative clauses.

2.2 Aspect

As noticed by Bertinetto (1991 [13]: 50), a special way to express the continuous aspect in Italian is the reiteration of the verbal form, as in example (3):

L’eroe cerca, cerca, cerca ma non trova nulla

‘The hero searches, searches, searches but does not find anything’
Bertinetto also notices that these forms are typical of the spoken language, in particular of fairy-tale style. Indeed, the repetition (for two, three or more times) of the same lexical material is a rather frequent strategy in spoken texts. This kind of repetition conveys a general meaning of intensification that, we may suppose, specialises according to the categorical nature of the repeated constituent. See for instance the example in (4) and its grid representation in Fig. 2.

(4) Ore 19.45 vado al pronto soccorso.. […] vado in radiologia, mi spoglio di nuovo […] torno al Pronto S. aspetto, aspetto, aspetto.. ma non ero un codice giallo??!!!!!
‘At 7.45 pm I go to the first aid.. […] I go to the radiology ward, I undress again […] I come back to the first aid. I wait, wait, wait.. but wasn’t it a yellow code??!!!!’

In this case the continuous aspect is solely expressed by the repetition of the verbal form aspetto (I) wait. One single occurrence of aspetto in the same contexts would simply denote a phase of the narration without stressing the continuity of the phase itself.

2.3 Reality status

The analyses of the Italian adverb magari (roughly corresponding to perhaps) carried out in Bonvino, Frascarelli and Pietrandrea (2007) [14], Pietrandrea (2007) [15] and Masini and Pietrandrea (2007) [16] have shown that this adverb has often scope on a constituent of different size and complexity (whether a clause, a verb, an argument or an adjunct) that belongs to a list, thus operating at both a clausal and an upper-clausal level. See the examples in (5) and (6) and their grid representations in Fig. 3 and Fig. 4, respectively.
(5) Che ne so poteva comparire una scenografia che che magari li riportava ne in un ambiente piuttosto che in una foresta in una giungla nel deserto
'I don’t know a set could appear that that maybe brought them back in a scene, or in a forest, in a jungle, in the desert'

(6) sono i mass media a sottolinearle con accenti drammatici e il tifo sconsiderato magari poco onesto
'Mass-media highlight them with dramatic accents and a thoughtless, maybe poorly honest, support'

The association of *magari* with this list-like TSD has been interpreted as an index of the non-factuality of this adverb. The fact that the speaker lists all the constituents in the scope of *magari*, placing them on the same level as possible options, indicates that he does not subscribe to any of them; rather he puts all of them forward as non-factual. Further, it was also found that the exact non-factual meaning of *magari* (potential, concessive, imperative, “non-factuality increase”) changes depending on the form taken by the list. Therefore, as a true construction in the technical sense, the TSD itself contributes to the overall meaning of the expression, which cannot be entirely captured by considering the adverb *magari* alone.
2.4 Approximation

The repetition of semantically related elements is one of the most common strategies to identify a referent in the spoken (and written, under certain circumstances) language. Whereas sometimes this repetition symbolises a lexical search for the “right word”, in other cases it aims at creating an “approximated” referent, which is clearly identifiable in the extralinguistic context, but cannot be fully linguistically categorized. Let us take example (7) and its grid representation in Fig. 5.

(7) poi torna sempre ‘sta mosca ‘sta zanzara quello che sia
then this fly this mosquito whatever it is is always coming back

Fig. 5. Grid representation of (7)

In (7), the postverbal subject is made up of three phrases or sequences (‘sta mosca ‘this fly’, ‘sta zanzara ‘this mosquito’, quello che sia ‘whatever it is’). These three sequences are not coordinated and are not endowed with an independent phrasal status; rather, they are the paradigmatic extension of a single grammatical position. In other words, the text in (7) does not present three subjects or three phrases with a subject function, but one complex subject that is built on the paradigmatic axis: we first have two semantically related (definite) nominal phrases and finally an expanded pronoun. The result is a well-identified non-generic referent (the sort of insect is there) with an approximated linguistic designation (it is a sort of insect). This semantic interpretation can only be caught if one takes into account the whole construction (i.e. the three constituents together) (cf. Bonvino 2005 [17]).

In (7) the approximation applies to the whole referent. However, the same approximating list-like construction also applies to single grammatical features. Consider example (8).

(8) è una che ci ha un figlio una figlia non so
‘it’s a woman who has a son a daughter I don’t know’

In this case, the approximating strategy has scope on the gender feature: the referent is clearly a child, but the speaker does not want to commit him/herself as regards his/her gender. In this way, we obtain the expression, in the topological
structure of discourse, of epicene gender, a grammatical value that cannot be expressed by Italian morphology.

In conclusion, whereas the first kind of approximation (cf. 7) can also be conveyed by other means in Italian, such as for instance the sorta/specie di ‘sort/kind of’ construction (cf. Masini 2007), the latter case can only be expressed by means of this particular TSD.

3 Computational applications

The project outlined in this contribution is only at its preliminary stage. We intend to further explore the hypotheses put forward in this paper by improving the methodology for detecting the TSDs, by building a fine typology of TSDs and, finally, by identifying the grammatical categories that can be expressed by the various TSDs and observing how the latter compete with constructions at other levels (such as syntactic or morphological constructions) for the expression of grammatical meaning.

This agenda is intended to address the challenge of offering new instruments for the practice of information retrieval. A thorough identification of the grammatical meanings associated to TSDs could, for example, provide the theoretical background for a finer semantic annotation of texts (especially spoken texts) which could ultimately end in a new strategy for information retrieval.

References