MACHINE-READABLE ENTAILMENTS WITH THE ITALIAN *PRENDERE* CONSTRUCTION EXPRESSING HITTING AND INSULTING EVENTS

Ignazio Mauro Mirto

Department of Culture e società, Università di Palermo, Italy

ABSTRACT

The Italian language features a little debated transitive construction with prendere *'to take/to catch' in which a prepositional phrase (PP) with an adverbial value occurs mandatorily (e.g.* Lui prese a pugni Leo *'He punched Leo'). Semantically, this construction often implies the use of physical force or verbal offence. In the hitting or insulting event, the notional subject generally is a [+ Human] Agent, whilst the notional direct object generally is a [+ Animate] Affectee ([1]: 4). It can be contended that* prendere, *which carries no literal meaning, is zero-valent and that the predicate assigning semantic roles is the PP. A computational tool will be illustrated, which automatically performs the following NLP / NLU tasks: it provides a reliable syntactic and semantic representation of the clause type, and it produces machine-readable entailments with three clause types, i.e. sentences with ordinary verbs, support verbs, and the causative verb* fare *'to make/to have'.*

KEYWORDS

Natural Language Processing, Recognizing Textual Entailment, Semantic Role Extraction, Adverbial PPs with a predicative value, Support verbs

1. INTRODUCTION

A phraseological unit (PU) of Italian involves one of the most frequent and polysemous verbs, i.e. *prendere* 'literally: take, catch, seize, pick up, etc.' (its closest synonym, i.e. *pigliare*, occurs less frequently), and a prepositional phrase (PP), as illustrated in (1) and (2) below (it is to be regretted that most examples of the *prendere* construction proposed below make reference to violence and this is certainly not the author's choice; the reason lies in the semantics of the construction, often related to injurious events):

(1)   Max ha preso a bacchettate   il compagno di classe
       Max has taken at stick-hitting.PL the classmate
       'Max hit his classmate with a stick'

(2)   I   senatori presero Cesare a pugnalate
       the senators took   Caesar at stabs
       'The senators stabbed Caesar'

To the best of our knowledge, no previous studies exist concerning the *prendere* construction exemplified above (the chapter in [2] dedicated to nouns ending in –*ata* contains no examples of this construction). Within its mandatory PP, 112 nouns can alternate with each other (this class is not restricted; 80 nouns (approximately 72%) are suffixed with –*ate*, as a single morpheme). The construction appears to have a collocational nature (cf. [3]: 596). Notice that *before*

combining with *prendere*, the PP is itself a case of *grammatical* collocation (see [4]). Thus there are grounds to consider the *prendere* construction as a case of *double* collocation. Its main characteristics are itemized below:

i. there is 'significant proximity' ([5]) between the collocating items (i.e. the PP and either *prendere* or *pigliare*);

ii. regarding the noun following the preposition: (a) it forms a binomial expression with the preposition *a* 'to / at'; (b) it must be bare, plural, and unmodified (with a few exceptions, e.g. *a male parole*, 'using bad words'); and (c) it designates a 'process' (as *bacchettate* 'stick-hitting' and *pugnalate* 'stabs' exemplify in (1) and (2));

iii. the verb *prendere* (which in Italian is often employed as a support verb, e.g. *Max ha preso una decisione* 'Max made a decision') contributes no meaning as regards semantic roles; and

iv. the PP is *not* a circumstantial (PPs are often optional in Italian); rather, it appears to function predicatively.

The purpose of this paper is to illustrate a computational tool, named NLP[ytaly] (cf. [6], [7] and the repository available on GitHub, https://github.com/ignaziomirto2017/nlpytaly/), which can automatically perform the following tasks: first, the identification of the PU (in the active, passive, and certain impersonal forms, e.g. the so-called *si* passivante); second, the extraction of the participants' semantic roles (expressed as *Cognate semantic roles*, see below); third, the identifying of the entailments which the PU establishes with simple sentences deploying two hypernyms: *colpire* and *insultare*; and last, the automatic identification of either one-way entailment or paraphrase relations with three distinct clause types *sharing the same content morphemes*: the ordinary verb construction, the support verb construction, and analytic causatives with the verb *fare* 'make / have / let' (*fare* causative construction).

2. WHY IS THE PP MANDATORY?

When PPs like those in (1) and (2) above combine with *prendere* (or *pigliare*), they cannot be removed without modifying the meaning of the sentence, as the comparison between (2) above and (3) below illustrates:

(3) I senatori presero Cesare
    'The senators took Caesar' (e.g. caught / kidnapped him)

However, with any other verb the PP is a genuine circumstantial (excluding *fare* 'make': e.g. *Fanno a cuscinate* 'They hit each other with pillows'), as the pair (4)-(5) demonstrates:

(4) I senatori uccisero Cesare a pugnalate
    'The senators killed Caesar by stabbing him (stabbed Caesar to death)'

(5) I senatori uccisero Cesare
    'The senators killed Caesar'

The very fact that the PP in (2) cannot be removed demonstrates that *a pugnalate* fulfills distinct functions in (2) and (4). Therefore, this PP is not a circumstantial and, not constituting a noun phrase, it cannot be considered as an argument of *prendere*. This raises a problem: what is its function? I believe that such a question should be conceived of as being intertwined with the non-literal use of the verb, in that, linguistically, nobody takes, catches, or seizes anything. In (3), *prendere* does license two arguments and it assigns to each a semantic role, whilst in (2) it evidently does not, otherwise the very same semantic roles would be identified. There are

therefore valid reasons for analyzing the PP as the genuine predicative core of the clause type and this is the rationale for it being unremovable. The same reasons also hold with the post-verbal noun in support verb constructions (e.g. *She shot him a scornful glance* vs. *She shot him*, see [6], [8], [9], [10]). That is, the verb *prendere* is to be interpreted as the indispensable element agreeing with the subject in person and number and carrying such features as e.g. tense and aspect. In addition, its frequent transitivity in literal uses makes available a syntactic grid for the subject and direct object functions. As far as the assignment of descriptive meaning is concerned, however, *prendere* appears to be inert.

3. MACHINE-READABLE ENTAILMENTS AND PARAPHRASES

This section describes one-way entailments and paraphrases which the *prendere* construction establishes with three clause types: the ordinary verb construction (OVC), the support verb construction (SVC), and the *fare* causative construction (FCC). The definitions of entailment and paraphrase on which I rely are those provided by [11]: 107-108 (notice that under this definition quantifiers can block the entailments, see [6]: 4-5).

3.1. The ordinary verb construction

Both the examples in (1) and (2) can be paraphrased with sentences constructed from ordinary verbs. Such paraphrases obtain by employing the verbs *bacchettare* 'hit with a stick' and *pugnalare* 'stab', respectively:

(6) Max ha bacchettato il compagno di classe
     'Max hit his classmate with a stick'

(7) I senatori pugnalarono Cesare
     'The senators stabbed Caesar'

By comparing (1) and (6), and ignoring the null-valent *prendere*, it can clearly be seen that the content morphemes in (6) are exactly the same as those found in (1) ('morphemic invariance', see [12]). Similarly, the content morphemes in (7) are also those which occur in (2). Moreover, the existence of a paraphrastic relation such as that between (1) and (6) means that these sentences entail each other, as sentences (2) and (7) also do.

The rationale for the above paraphrases is rather simple: the noun within the PP in e.g. (1) does have a verbal counterpart *with the same content morpheme*. That is, the noun-verb pair *bacchettate* : *bacchettare* exists in Italian. Overall, there are approximately 20 such noun-verb pairs (which is 17% of all nouns occurring in the clause type), and they include diminutives (e.g. *schiaffi / schiaffetti* 'slaps / little slaps') and augmentatives (e.g. *spinte / spintoni* 'pushes / big pushes'). It is of use to refer to this first class in the data structure of the tool as *Group A*. 14 other nouns (12.5%) do not have a verbal counterpart, but one can be easily located, due to their synonymic relationship with a number of nouns in *Group A*, e.g. *ceffoni* 'big slaps' (*ceffonare* does not exist), to be associated with *schiaffi* 'slaps' and then paired with the cognate verb *schiaffeggiare* 'to slap'. This latter list will be referred to as *Group B*.

Again, no verbal counterpart with the same content morpheme is available with 69 other nouns (61.5%, *Group C*). However, with such nouns an entailment relationship can be straightforwardly constructed. This is by means of sentences in which the verb *colpire* 'hit' occurs (this association also holds true with *Groups A* and *B*), as the pair formed by (8) and (9) exemplifies. Notice that most PPs formed with the nouns in *Groups A*, *B* and *C* can be replaced with the phrase '*a colpi di* + noun', with no changes in meaning (cf. [13]). For example, the sentences *Lo ha preso a gomitate / a bottigliate* 'He hit him with his elbow / with a bottle'

respectively bear the same meanings as *Lo ha preso a colpi di gomito*, *Lo ha preso a colpi di bottiglia*.

(8) Loro presero l'uomo a cuscinate
   they took   the man at blows.with.a.pillow
   'They beat the man with a pillow'

(9) Loro hanno colpito l'uomo (con un cuscino)
    'They beat the man (with a pillow)'

Finally, the remaining 10 nouns (9%; only two of them have a verbal counterpart, as in *Group A*) call to mind cases of *verbal offence*. With such nouns, the counterpart constructed with an ordinary verb can regularly be rendered with *insultare* 'to insult':

(10)   Loro presero l'uomo a   parolacce / insulti
          they took   the man at bad.words / insults
          'They addressed the man using bad words'

(11)   Loro insultarono l'uomo
          'They insulted the man'

These nouns can be labelled as *Group D*. It is worth noticing that with pairs such as (10) and (11) a one-way entailment is observed rather than a mutual entailment: if (10) is true, then (11) must be true. It is also evident that the opposite does not hold, in that someone can be insulted in a number of ways.

3.2. The support verb construction

Mutual entailment relationships can be also established between such sentences as the following:

(12)   Max prese a morsi la mela
          Max took at bites the apple
          'Max bit the apple'

(13)   Max diede dei morsi alla mela
          Max gave some bites at the apple
          'Max bit the apple'

Sentence (12) differs from the above examples of the construction under scrutiny only on account of its direct object (i.e. *la mela* 'the apple') being [- Animate]. On the other hand, sentence (13) is an instance of the well-known support verb construction (SVC) (see [8]). As the translation in (13) clearly indicates, the verb *diede* 'gave' possesses a blank valence and it carries no literal meaning (see [14]).

Let us now focus on the intersection between the data structure which is currently available in NLP[ytaly] for the SVC (approximately 770 predicate nouns such as *morso* 'bite' in (13)) and the four groups in the data structure which were identified in the preceding Section. All the nouns in *Group A* can also be employed as the predicate noun of a SVC, whilst only 8 nouns in *Group B* have a counterpart in the SVC. As for *Group C*, approximately 40 of the nouns included can also be employed in the SVC, whilst only four of those in *Group D* (*insulti* 'insults', *minacce* 'threats', *offese* 'offences', and *pernacchie* 'raspberries') are included in the SVC. Of

importance, the entailment between such sentences as (12) and (13) is reciprocal: the truth of the former guarantees the truth of the latter, and *vice versa*.

3.3. The *fare* causative construction

Unlike in Subsections 3.1 and 3.2, the entailments between an FCC and the *prendere* construction cannot be mutual, and for compelling reasons. Let us consider an FCC sentence such as *Leo fece schiaffeggiare il ragazzo da Max* 'Leo had Max slap the boy'. The truth of this sentence guarantees the truth of at least two other sentences (with which an entailment relationship is therefore established), as illustrated in (14) and (15) below. The former is an additional instance of the construction with *prendere*, whilst the latter is another FCC in which the cause of the slapping event is the subject *Leo*:

(14)   Max prese a schiaffi il ragazzo
         Max took at slaps   the boy
         'Max slapped the boy'

(15)   Leo fece prendere il ragazzo a schiaffi da Max
         Leo made take   the boy   at slaps   by Max
         'Leo had Max slap the boy'

An entailment relationship also characterizes (14) and (15), albeit in one direction only: (15) entails (14), whilst (14) does not entail (15). As suggested above, the reason is simple: if compared to (14), the FCC in (15) features an extra participant (and therefore an additional semantic role), i.e. the causer *Leo* (namely the sentence subject, licensed by *fare*, cf. [14]) and this precludes any reciprocal entailment.

4. GENERAL PICTURE, SEMANTIC ROLE EXTRACTION, AND CONCLUDING REMARKS

The network of entailments and paraphrases described in Section 3 can be represented diagrammatically as in Figure 1:
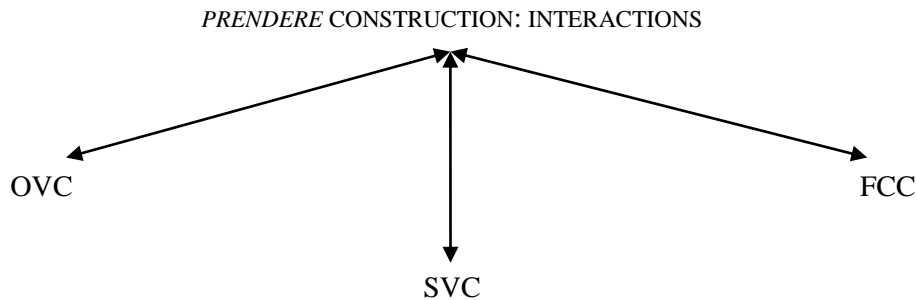

Figure 1: Machine-readable entailments obtainable with the tool NLP[ytaly]

The arrows illustrate the relationships which the *prendere* construction establishes with the SVC (exemplified with pair (12)-(13)), the FCC (see pair (14)-(15)), and the OVC (see pairs (1)-(6), (2)-(7), (8)-(9), and (10)-(11)).

The meaning which NLP[ytaly] automatically extracts is based on a novel type of semantic role, labelled ***Cognate Semantic Role*** (CSR, see [6], [7], [9], [10]). For example, with an OVC sentence such as (7) (*I senatori pugnalarono Cesare*), the NLP[ytaly] tool extracts the CSRs >the one who stabs< and >the one who is stabbed< (these semantic roles are expressed in Italian).

These cognate semantic roles are subsequently paired with the subject and direct object, respectively. The outcome will therefore be two-fold: (a) 'the senators are those who stab', and (b) 'Caesar is he who is stabbed'. Sentence (2) (*I senatori presero Cesare a pugnalate*) is initially identified as a *prendere* construction: *prendere* is thus ignored and CSRs obtain from the PP having recourse to the verb which is cognate to *pugnalate* 'stabs', i.e. *pugnalare* 'to stab'. Thus, the same pairings obtained from (7) are also obtained from (2) and the entailment between (2) and (7) becomes machine-readable, inasmuch as the set of pairings extracted from the former sentence will be identical to that of the latter. This also holds true for the relationship between the *prendere* construction and the SVC, whilst a one-way entailment (thus devoid of a paraphrase) obtains with the FCC.

It is worth noting that no approach which centers on the verb as the only possible assigner of semantic roles can attain the above-mentioned entailments and paraphrases. The rationale for this is that, in the *prendere* construction (as well as in the SVC), the verb plays no role in the assignment of cognate semantic roles.

REFERENCES

[1] Fillmore, J. Charles (1968), "The case for case", in Bach, E. & Harms, R.T. (eds.) *Universals in linguistic theory.* New York, Holt, Rinehart & Winston, 1-88.

[2] Gaeta, Livio (2002), *Quando i verbi compaiono come nomi. Un saggio di Morfologia Naturale*, Milano, FrancoAngeli.

[3] Krishnamurthy, Ramesh (2006), "Collocations", in *Encyclopedia of Language and Linguistics* (2nd Edition), R. Brown (Editor-in-chief), Vol. 2, Amsterdam, Elsevier Science, 596-600.

[4] Faloppa, Federico (2010). "Collocazioni", in *Enciclopedia dell'italiano Treccani*. Available online: https://www.treccani.it/enciclopedia/collocazioni_(Enciclopedia-dell'Italiano)/.

[5] Halliday, Michael A. K. & Ruqaiya Hasan (1976), *Cohesion in English*. London – New York, Hove Longman.

[6] Mirto, Ignazio Mauro (2021), "Automatic extraction of semantic roles in support verb constructions", *International Journal on Natural Language Computing* (IJNLC) Vol.10, No. 3, 1-10.

[7] Mirto, Ignazio Mauro (2022), "Measuring meaning", in Arai, K. (ed.), *Intelligent Computing*, LNNS 283, 1054-1067.

[8] Gross, M. (1981), "Les bases empiriques de la notion de prédicat sémantique", *Langages*, vol. 63, 7-52.

[9] Mirto, Ignazio Mauro (2019), "The role of cognate semantic roles: Machine translation support for support verb constructions", Paper presented at the 13th International Conference NooJ 2019, Hammamet, Tunisia, June 7-9. Available online: (PDF) The role of cognate semantic roles: Machine translation support for support verb constructions | Ignazio Mauro Mirto - Academia.edu.

[10] Mirto, Ignazio Mauro (2020), "Natural Language Inference in Ordinary and Support Verb Constructions", in Dong *et al* (eds.), *Distributed Computing and Artificial Intelligence,17th International Conference,* Springer Nature Switzerland, 124-133.

[11] Hurford, James R. & Brendan Heasley (1983). *Semantics: A coursebook*. Cambridge, Cambridge University Press.

[12] Harris, Zellig S. (1981), *Papers on Syntax*, H. Hiż (eds.), London, D. Reidel Publishing Company.

[13] Gross, Gaston (1984), "Étude syntaxique de deux emplois du mot 'coup'", in *Lingvisticae Investigationes*, vol. 8: 37-61.

[14]   La Fauci, Nunzio & Ignazio M. Mirto (2003), Fare. *Elementi di sintassi*. Pisa, ETS.

Author

Ignazio Mauro Mirto (Ph.D. Cornell University) is Associate Professor of General Linguistics at the University of Palermo. His current work encompasses the area of morphosyntax, information structure, and determiners as inflectional morphemes for the syntactic closure of noun phrases functioning as arguments.