Encoding semantic phenomena in verb-argument combinations

Elisabetta Jezek[1], Costanza Marini[1,2], Emma Romani[1]

[1]University of Pavia, Dipartimento di Studi Umanistici, Strada Nuova 65, 27100 Pavia
[2]University of Bergamo, Dipartimento di Lingue, Letterature e Culture Moderne, via Salvecchio 19, 24129 Bergamo
E-mail: e.jezek@unipv.it, costanza.marini01@universitadipavia.it, emma.romani01@universitadipavia.it

Abstract

In this paper, we report the classification we adopted in two electronic resources of corpus-derived verbal patterns for Italian and Croatian (T-PAS and CROATPAS) to account for three different semantic phenomena that we observed occurring between nouns and verbs in valency structure contexts: Semantic Type alternation, Semantic Type shift (metonymy), and Complex Type exploitation. After presenting the two resources in the context of similar projects (Section 2), in Sections 3, 4, and 5 we examine the three phenomena in detail and show how we registered them in the editor we developed for this purpose, called Skema. The encoding of these phenomena in the editor is of paramount importance for being able to query them in the interface of the two resources, which will soon be publicly available online. In Section 5, we draw our conclusions and suggest possible ways to use the annotated data.

Keywords: pattern resource; verb argument structure; semantic type; corpus analysis; word sense

1. Introduction

Lexical resources traditionally rely on lists of word senses, although several studies have long shown that word senses are very slippery entities (Kilgarriff, 1993), and that sense inventories fail to capture the large spectrum of meanings words acquire in their context of use. From a theoretical perspective, the variation in the senses of a word stems from the fact that natural languages are semantically flexible, that is, the meaning of a word varies from occurrence to occurrence as a function of the interaction with the other words it combines with, and with the context of utterance (Pustejovsky, 1995; Recanati, 2002). Within this framework, in lexicography word senses are then better conceived as abstractions from clusters of corpus citations (Kilgarriff, 1993: 91).

In this paper, we present two resources of verbal patterns that take this background into account, and address the problem of encoding the sense variation that can be observed in the nouns filling the argument positions in the pattern, which we assume are triggered by the verb the nouns combine with. Specifically, we report the classification we adopted in two inventories of predicate-argument structures – namely, T-PAS for Italian (Ježek et al., 2014) and its sister project CROATPAS for Croatian (Marini & Ježek, 2019) – to account for three different semantic phenomena that may affect nouns within a valency structure context: Semantic Type Alternation, Semantic Type Shift (Metonymy), and Complex Type Exploitation. This is possible thanks to a shared System of Semantic Types used to classify the semantics of arguments (Ježek, 2019), to the compositional principles of type coercion and type exploitation inspired by the Generative Lexicon (Pustejovsky & Ježek, 2008), to the methodological framework of corpus analysis adopted from Hanks (2013), and, last but not least, thanks to the editor that was developed to encode the phenomena at play (Baisa et al., 2020).

The structure of the paper is as follows: in Section 2 we introduce the two resources; in Section 3 we provide examples of Semantic Type Alternation occurring in different syntactic positions; in Section 4 we discuss Metonymy; in Section 5 we illustrate Complex

Type Exploitation. Finally, in Section 6 we highlight the usefulness of encoding these phenomena in electronic resources.

2. The resources: T-PAS and CROATPAS

T-PAS (Ježek et al., 2014) and CROATPAS (Marini & Ježek, 2019) are two corpus-derived resources consisting of repositories of Typed Predicate-Argument Structures for Italian (T-PAS) and Croatian (CROATPAS) verbs. Both projects are being developed at the University of Pavia with the technical support of *Lexical Computing Ltd.* and are intended to be used for linguistic analysis, language teaching, and computational applications. The resources share their organisation as regards four fundamental components:

1. a repository of corpus-derived predicate argument structures (called *patterns*) with semantic specification of their argument slots, e.g. [Human] drinks [Beverage];
2. an inventory of ca. 200 corpus-derived semantic classes (called *Semantic Types*) organised in a hierarchy (called *System of Semantic Types*), used for the semantic specification of the arguments;
3. a corpus of annotated sentences that instantiate the different patterns of the verbs in the inventory. Corpus lines are tagged with their respective pattern numbers and anchored to the verb they feature, which is the lexical unit of analysis;[1]
4. an editing system called Skema (Baisa et al., 2020), which allows the registration of patterns and all the syntactic and semantic information associated therewith, and facilitates the manual annotation of corpus instances (directly linked to the patterns).[2]

Typed predicate-argument structures are patterns that display the semantic properties of verbs: for each meaning of a verb, a specific pattern is provided. As referenced above, the patterns are corpus-derived, i.e. they are acquired through the manual clustering and annotation of corpus instances, following the CPA methodology (Hanks, 2013). Currently, T-PAS contains 1160 implemented verbs, 5,529 patterns, and ca. 200,000 annotated corpus instances, while CROATPAS contains 180 verb entries, 683 patterns and ca. 23,000 annotated corpus lines.

In the resources, each pattern is labelled with a pattern number and connected to a list of corpus instances realising that specific verb meaning. The Skema editor (see Figure 1) enables the registration of different semantic and lexical information in each pattern, more specifically:

1. the *verb*, which in T-PAS is generally in its infinitive form - e.g. *bere* (Eng., 'to drink');

---

[1] The reference corpora for the resources are two web corpora, namely *ItWac* (reduced) for T-PAS and *hrWac 2.2* for CROATPAS. *ItWac* (reduced) contains around 935 million tokens, while *hrWac 2.2* contains roughly 1.2 billion tokens.

[2] Skema (Baisa et al., 2020) is a corpus pattern editor system implemented to facilitate the management of manual annotation of concordance lines with user-defined labels and the editing of the corresponding patterns in terms of slots, attributes and other features following the lexicographic technique of CPA (Hanks, 2013).

2. the *Semantic Types* (e.g. [Human], [Beverage], always portrayed within square brackets), specifying the semantics of the arguments selected by the verb. Semantic Types can be found on six different arguments positions: *subjects* (portrayed in red), *direct objects* (green), *adverbials* (grey), *clausals* (violet), *predicative complements* (blue), *prepositional complements* (orange, only in T-PAS), and *indirect complements* (light blue, only in CROATPAS).

3. the *sense description*, i.e. a brief definition of the meaning of the verb in that specific pattern, which usually features the same Semantic Types registered in the pattern in question;

4. a *lexical set* (optional) for each Semantic Type in the pattern, i.e. a selection of the most representative lexical items instantiating that Semantic Type (e.g. *vino* = 'wine' | *birra* = 'beer' | *aranciata* = 'orange juice' are good candidates for the lexical set of [Beverage]);[3]

5. the *roles* (optional) played by some specific Semantic Types in certain contexts: in particular, the Semantic Type [Human] can acquire the role of Athlete, Doctor, Musician, Host, Guest, Writer, etc., depending on the verb selecting it as an argument;[4]

6. the *features* (optional) associated with the Semantic Types, i.e. certain semantic characteristics required by the pattern syntax (e.g. Plural) or by the specific verb meaning (e.g. Female, Negative, Visible);[5]

7. *prepositions* (for prepositional and indirect complements), *particles* (for adverbials), *complementisers* (for clausals), *quantifiers*, and *determiners* (for lexical sets), which can be implemented according to the specific argument position in question.

The System of Semantic Types used to classify the semantics of arguments (Pustejovsky et al., 2004; Ježek, 2019) is a hierarchy of general semantic categories obtained by manual clustering of the lexical items found in the argument positions of corpus-derived valency structures. The System currently contains ca. 200 Semantic Types that are hierarchically organised on the basis of the 'is a' (subsumption) relation (e.g. [Human] is an [Animate]).[6] The System of Semantic Types is shared by both resources.

Figure 1 shows the general organisation of both resources in the Skema editor (using patterns and corpus examples from the Italian T-PAS resource) with its four components used by the annotators to compile the patterns:[7]

---

[3] Lexical sets appear next to their respective Semantic Types, in curly brackets.

[4] In Skema, Roles appear within square brackets, next to the Semantic Types they apply to, and preceded by '=', e.g. [Human = Doctor].

[5] In Skema, Features appear within square brackets, after the Semantic Types they apply to, and preceded by ':', e.g. [Human : Plural].

[6] The System of Semantic Types, together with definitions and examples for each Type, is made accessible to lexicographers through a customised function of Skema, so that it can be readily consulted while editing the patterns.

[7] The Skema editor is only accessible to the annotators working on the projects; the online public version based on Skema will display the patterns in a graphical interface that can be browsed.

Figure 1: The general structure of the resources (based on T-PAS) with the four main components as encoded in the Skema pattern editor (from the top of the image): patterns, pattern editor, corpus, System of Semantic Types

When it comes to pattern resources, it is necessary to mention some noteworthy projects revolving around several different languages. Chronologically, the first project where Corpus Pattern Analysis was applied was the *Pattern Dictionary of English Verbs* (PDEV) (Hanks & Pustejovsky, 2005), which is being developed at the Research Institute for Information and Language Processing of the University of Wolverhampton. An equivalent Spanish project is *Verbario* (Renau & Nazar, 2021), developed at the Pontifical Catholic University of Valparaíso (Chile). As for Dutch, a recent tool combining verb patterns, collocations and idioms is *Woordcombinaties* (Colman & Tiberius, 2018), which is being developed in Leiden at the *Instituut voor de Nederlandse taal*. Last but not least, another Italian pattern dictionary is currently being designed at the University of Heidelberg (Germany). The project is aimed at creating a learner's dictionary with phraseological disambiguators (Di Muccio-Failla & Giacomini, 2017).

In the rest of the paper, we will focus on the encoding of the three semantic phenomena that we have detected while building the pattern resources, and encoded in Skema. They are: Semantic Type Alternation, Semantic Type Shift (Metonymy), and Complex Type Exploitation.

3. Semantic Type Alternation

Let us start with the most frequent phenomenon, Semantic Type Alternation. When different Semantic Types alternate on the same argument slot within the same verb sense – i.e. within the same *pattern* – a Semantic Type alternation is at play. Semantic Type alternations are a pervasive phenomenon in both the T-PAS and CROATPAS resources and are graphically encoded by adding vertical bars "|" (which stand for the OR operator) between the alternating Semantic Types.

An example of Semantic Type Alternation on the subject position is the one between [Human] and [Wind] in the context of pattern 1 of the Italian verb *rimuovere* 'to remove' (Figure 2).
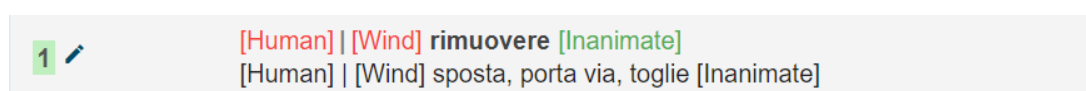

Figure 2: Pattern 1 of the Italian verb *rimuovere* 'to remove'

The following corpus lines (Figure 3) can be considered to be instantiations of the pattern:

Figure 3: Corpus lines linked to pattern 1 of the Italian verb *rimuovere* 'to remove' with subjects in red

Let us compare the two highlighted sentences: *Il <u>sindaco</u> di Pieve ha fatto rimuovere un grande striscione*, 'The <u>major</u> of Pieve had a big banner removed', in which the word

sindaco is an instance of the Semantic Type [Human], and *Il vento e l'acqua potrebbero rimuovere la polvere di Uranio impoverito dalla superficie del veicolo*, 'Wind and water may remove uranium dust from the vehicle's surface', in which *vento* instantiates the Semantic Type [Wind]. In both cases, the meaning of the verb is the same, that is, 'removing' something. In this case, the two Semantic Types are not linked by any kind of relation. This is not true for all Semantic Type Alternations, as we will show below.

Turning now to the object position, an interesting alternation taking place on the object slot of pattern 3 of the Croatian verb *otkriti* 'to reveal' is [Part of Body | Body] (Figure 4).[8]

Figure 4: Pattern 1 of the Croatian verb *otkriti* 'to reveal'

Unlike the previous case, in this case the two alternating Semantic Types are clearly linked by a meronymic relationship of Part/Whole. For this reason, it is all the more obvious that their alternation does not imply any meaning shift in the verb, as is testified by the highlighted sentences from Figure 5: *(Korzet) je otkrio njezina gola ramena*, 'The corset revealed her bare shoulders', and *Skinula je glamuroznu haljinu i preodjenula se u žuti bikini, koji je otkrio na baš savršeno tijelo*, 'She took off the glamourous dress and changed into a yellow bikini, which revealed a truly perfect body'.

Figure 5: Corpus lines linked to pattern 3 of the Croatian verb *otkriti* 'to reveal', with objects in green

To provide an idea of the frequency of Semantic Type alternations, we report some raw figures from T-PAS. For each argument position (column 1), we provide the number of patterns that include that argument slot in their valency structure (column 2) and the number of patterns featuring at least one Semantic Type Alternation in that position (column 3).

The final line of Table 1 displays the overall number of T-PAS patterns (column 2) and the overall number of T-PAS patterns with at least one alternation on any argument position (column 3). Note that these numbers are lower than the sum of the elements in each column, since the same pattern can encompass more than one argument slot (e.g. a subject and an object), each potentially bearing a Semantic Type Alternation. However, we can still state that nearly 45 percent (2,468 out of 5,529) of the patterns

---

[8] Since Croatian is a Slavic language equipped with its own case system to express the relationships between sentence components, the Croatian version of the Skema editor has been enriched by adding explicit bottom-right case markings on each argument slot, such as *nominative* or *accusative.*

Argument	No. of patterns	No. of patterns with Semantic Type Alternation
Subject	5,503	1,687
Object	3,184	1,097
Prepositional complement	1,668	450
Adverbial	379	0
Clausal	435	9
Predicative complement	108	16
Overall	5,529	2,468

Table 1: T-PAS patterns featuring Semantic Type Alternations for each argument position

in the inventory feature a Semantic Type alternation on at least one of their argument positions.

4. Semantic Type Shift

In both T-PAS and CROATPAS, the changes in meaning of an argument caused by metonymic displacements are not encoded as Semantic Type Alternations but as Semantic Type Shifts. Following Pustejovsky (1995), we assumed that such shifts take place when a Semantic Type is forced by the verb to be understood as a different one (which satisfies its semantic selectional requirements or preferences).

Three clear-cut cases of metonymy are offered by the sentence *Ho letto Dante, Moravia, Calvino*, 'I have read Dante, Moravia, Calvino, ' from Figure 6, where the Italian verb *leggere*, 'to read', triggers a shift from [Human = Writer] to [Document]. Unlike in the first highlighted sentence - *Ho ultimamente letto il libro di Harry Potter*, 'I have recently read the Harry Potter book' – each time the verb *leggere* combines with the name of an author on the object position, the well-known Author/Work metonymy takes place, forcing that person to be interpreted as the *document he or she has written.*

Figure 6: Corpus lines linked to pattern 2 and subpattern 2.m of the Italian verb *leggere* 'to read'

As shown in Figure 7, the metonymy at play is encoded in Skema as a sub-pattern of the main pattern [Human] reads [Document] (Romani & Ježek, 2020; Marini & Ježek, 2020). Note that the labels of the subpattern and of the metonymic corpus lines linked to it are the same: they start with the same number as the main pattern label and end in '.m', which stands for *metonymic.*

Figure 7: Pattern 2 and metonymic subpattern 2.m of the Italian verb *leggere* 'to read'

Let us now consider the Semantic Type shift taking place in the last three corpus instances from Figure 8 – one of them being *Studirao je violončelo*, 'He studied cello' – and compare them to the first three non-metonymic examples – e.g. *Studirala si komparativnu književnost i povijest umjetnosti*, 'You studied comparative literature and art history'.

Figure 8: Corpus lines linked to pattern 1 and subpattern 1.m of the Croatian verb *studirati* 'to study'

Indeed, having studied comparative literature and art history implies having acquired a deep knowledge of those theoretical fields, whereas having studied a [Musical Instrument] means "having studied how to play it". This last piece of information is not explicitly stated, but is the result of a metonymic shift triggered by the verb *studirati*, 'to study', which requires either a theoretical [Field of Interest] or an [Activity] in the direct object slot (Figure 9), thus forcing [Musical Instrument] to be understood as the [Activity] of playing it.

Figure 9: Pattern 1 and metonymic subpattern 1.1.m of the Croatian verb *studirati*, 'to study'.

5. Complex Type Exploitation

In our System of Semantic Types, we acknowledge the existence of Complex Types. Complex Types are unique Semantic Types "made up" of two (or more) components (Pustejovsky & Ježek, 2008): for example, [Institution] is a Complex Type made up of

[Abstract Entity] and [Human Group]. In the Skema editor, we currently encode Complex Types as "simple" Semantic Types (e.g. [Institution]). However, we keep track of their internal complexity by locating them in multiple places in the System of Semantic Types, as sub-types of their components: for example, as one can see in Figure 10, the Complex Type [Institution] is located both under [Human Group], which is a kind of [Animate] entity, and under [Abstract Entity].[9] We call this phenomenon multiple inheritance, meaning that a Complex Semantic Type inherits from different Types of the hierarchy.
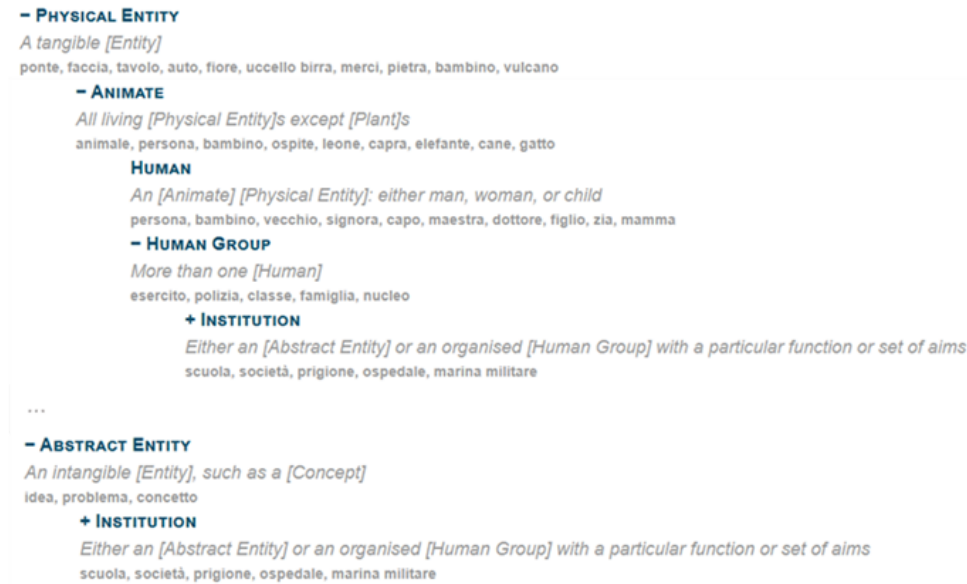

Figure 10: The System of Semantic Types used in T-PAS and CROATPAS, where the Complex Type [Institution] is registered both as a type of [Human Group] and [Abstract Entity]

That having been said, we encode a Complex Type Exploitation when a verb exploits only one of the components of a Complex Semantic Type associated with an argument. In this case, no metonymy occurs because there is no change of referent, as is the case in the examples in Section 4. In the following, we provide some examples of Complex Type Exploitation in the patterns of T-PAS and CROATPAS, focusing on two Complex Types, and highlighting which component is exploited. We also include instances of co-predication, i.e. contexts in which both components are simultaneously exploited.[10]

The first Complex Type we examine is [Institution], whose components are [Abstract Entity] and [Human Group]. In pattern 5 of the T-PAS verb *accettare*, 'to accept', for example, the verb only exploits the human component of the Complex Type [Institution] of its subject (Figure 11), as the act of accepting someone is typical of a [Human] or a [Human Group]:[11]

---

[9] Each component of a Complex Type is a "real" Semantic Type, which can also be used independently of the Complex Type.

[10] Recall that co-predication is the test traditionally used in linguistic and ontological studies to identify Complex Types (Pustejovsky, 1995).

[11] Even though the Semantic Types [Human] and [Human Group] are connected by the Whole/Part relationship (given that a [Human Group] is a group of more than one [Human]), they are not related in our System as the only relation that we consider is the relation of subsumption, e.g. 'is-a-type-of'. They are listed at the same level and subtypes of [Animate]s
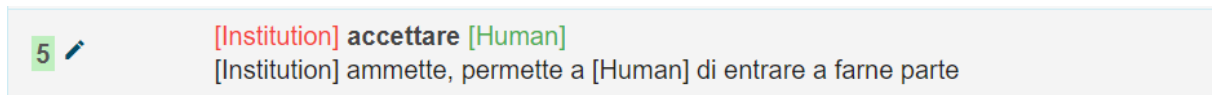
Figure 11: Pattern 5 of the Italian verb *accettare*, 'to accept', featuring the Complex Type [Institution] in the subject position, exploited in its [Human Group] component.

This pattern is instantiated by corpus examples such as *Non tutte le università accettavano le donne e l'Università di Varsavia era tra queste* 'Not all universities accepted women, and the University of Warsaw was one of them', as highlighted in Figure 12.

Figure 12: Corpus instances for the verb *accettare* 'to accept' and instantiations of the Semantic Type [Institution]

Conversely, in pattern 7 from Figure 13, the verb *dissolvere*, 'to dissolve', only selects the [Abstract Entity] component of [Institution]. The meaning of the verb in this case is figurative:
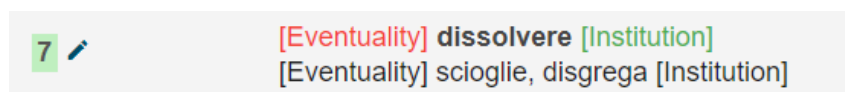

Figure 13: Pattern 7 of the verb *dissolvere*, 'to dissolve', featuring the Complex Type [Institution] in the object position, exploited in its [Abstract Entity] component

An example of this kind of exploitation is *Le invasioni barbariche dissolvono l'Impero* 'Barbaric invasions disintegrate the Empire', as highlighted in Figure 14:

Figure 14: Corpus instances of the verb *dissolvere*, 'to dissolve', and instantiations of the Semantic Type [Institution]

We may also consider corpus sentences that display co-predication, that is, cases in which both components are exploited with regard to the same argument, as in *L'Università di Padova fu fondata nel 1222 ed è stata la prima al mondo ad accettare studenti ebrei*, 'The University of Padova was founded in 1222, and it was the first in the world to accept Jewish students.' In this case, the verb *fondare*, 'to found', taking [Institution] as an object, exploits the [Abstract Entity] component, whereas the verb *accettare*, as in the previous sentence, exploits the [Human Group] component.

As a second case, we consider examples of the exploitation of the Complex Type [Artwork], whose components are [Artifact] and [Concept]. For example, the Croatian verb *izlagati*, 'to exhibit', exploits only the Artifact component of this Complex Type, which we can find in the object position of pattern 1 in Figure 15 from CROATPAS.

Figure 15: Pattern 1 of the Croatian verb *izlagati*, 'to exhibit', and the Complex Type [Artwork] in the object position, exploited in its [Artifact] component

After all – as we can see from sentences such as *Predstavljeni su dizajneri koji će ove sezone izlagati svoje kreacije*, 'The designers that will exhibit their creations this season have been presented', from Figure 16 – artistic creations need to be physical entities in order to be exhibited.

Figure 16: Corpus lines linked to pattern 1 of the Croatian verb *izlagati*, 'to exhibit'.

Conversely, we can consider pattern 2 of the Italian verb *partorire*, 'to give birth', for the exploitation of the [Concept] component of the Complex Type [Artwork] (Figure 17). Note that *partorire* does not select the physical component of [Artwork], since its meaning is figurative: one cannot literally 'give birth to an [Artwork]', but rather we can talk of conceiving it in our mind, which is why we are only exploiting its conceptual component.

Figure 17: Pattern 2 of the Italian verb *partorire*, 'to give birth', featuring the Complex Type [Artwork] in the object position, exploited in its [Concept] component

As we can see from the following corpus instances (Figure 18), the meaning is clearly 'to mentally conceive something', as in the sentence *Il grande Kenji Inafune ha finalmente partorito il suo nuovo capolavoro*, 'The great Kenji Inafune has finally given birth to its new masterpiece'.

Figure 18: Corpus instances for *partorire*, 'to give birth', and instantiations of the [Artwork] Semantic Type

Finally, consider this instance of co-predication in which both components ([Concept] with *partorire*, 'to give birth' and [Artifact] with *presentare*, 'to present') are present: *Nel 1501 Leonardo da Vinci partorì un'opera di straordinaria importanza, che fu presentata al sultano Bezajet II: si trattava di un ponte …* (Eng.: 'In 1501 Leonardo da Vinci gave birth to an artwork of extraordinary importance, which was presented to Sultan Bezajet II: it was a bridge …').

6. Conclusions and future prospects

In this paper, we have shown how the semantic phenomena that take place in verb-argument combinations are encoded in two electronic resources dedicated to the description of corpus-derived verb-argument structures present in Italian and Croatian. In particular, we have discussed cases in which there is an alternation of Semantic Types on the same argument position within the same verb sense, cases where there is a Semantic Type Shift, and, finally, cases where a single component of a Complex Type denoted by a noun is exploited in the composition. We have shown how these data are currently stored in the off-line editor that we developed, called Skema.

In the near future, the data will be made public through a graphical interface, where users will be able to search for the three phenomena by browsing for the pattern and subpattern inventory (accompanied by Good Dictionary EXamples selected from the annotated corpus for each pattern (Kilgarriff et al., 2008)). Users will also be able to query the system of Semantic Types and the different argument positions (subject, object, prepositional complement, indirect complement, clausal, and predicative complement), both one at a time, as well as in combination.

The data in the two resources can be useful for linguistic research in syntax and semantics, for example, for studies aiming at classifying verbs based on the semantic selection of their arguments. Moreover, they can be useful for corpus-based approaches to language teaching, and possibly as a gold standard in natural language processing tasks involving figurative language recognition in accordance with Shutova et al. (2013), who used a

combination of corpus analysis and knowledge base extraction to predict classes of words in order to identify instances of logical metonymy.

7. Acknowledgments

8. References

Baisa, V., Tiberius, C., Ježek, E., Colman, L., Marini, C. & Romani, E. (2020). Skema: A New Tool for Corpus-driven Lexicography. In *Proceedings of the 19th EURALEX International Congress.*

Colman, L. & Tiberius, C. (2018). A Good Match: a Dutch Collocation, Idiom and Pattern Dictionary Combined. In *Proceedings of the 18th EURALEX International Congress.* Ljubljana, Slovenia.

Di Muccio-Failla, P. & Giacomini, L. (2017). Designing a Learner's Dictionary with Phraseological Disambiguators. In *Proceedings of the 2nd International Conference on Computational and Corpus-Based Phraseology (Europhras).* London, UK.

Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations.* Cambridge: The MIT Press.

Hanks, P. & Pustejovsky, J. (2005). A Pattern Dictionary for Natural Language Processing. *Revue Française de Linguistique Appliquée*, 11(2), pp. 63–82.

Ježek, E. (2019). Sweetening Ontologies Cont'd: Aligning bottom-up with top-down ontologies. In A. Barton, S. Seppälä & D. Porello (eds.) *Proceedings of the Joint Ontology Workshops 2019.* Graz, Austria.

Ježek, E., Magnini, B., Feltracco, A., Bianchini, A. & Popescu, O. (2014). T-PAS: A resource of corpus-derived Types Predicate-Argument Structures for linguistic analysis and semantic processing. In *Proceedings of LREC.* pp. 890–895.

Kilgarriff, A. (1993). I Don't Believe in Word Senses. *Computers and the Humanities*, 31(2), pp. 91–113.

Kilgarriff, A., Husák, M., Mcadam, K., Rundell, M. & Rychlý, P. (2008). GDEX: automatically finding good dictionary examples in a corpus. In *Proceedings of the 13th EURALEX International Congress.* Barcelona, Spain.

Marini, C. & Ježek, E. (2019). CROATPAS: A Resource of Corpus-derived Typed Predicate Argument Structures for Croatian. In *Proceedings of the 6th Italian Conference on Computational Linguistics.* Bari, Italy.

Marini, C. & Ježek, E. (2020). Annotating Croatian Semantic Type Coercions in CROATPAS. In *Proceedings of the 16th Joint ACL-ISO Workshop on Interoperable Semantic Annotation.* Marseille, France.

Pustejovsky, J. (1995). *The Generative Lexicon.* Cambridge, MA: MIT Press.

Pustejovsky, J., Hanks, P. & Rumshisky, A. (2004). Automated Induction of Sense in Context. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING).* Geneva, Switzerland.

Pustejovsky, J. & Ježek, E. (2008). Semantic Coercion in Language: Beyond Distributional Analysis. *Italian Journal of Linguistics/ Rivista Italiana di Linguistica*, 20(1), pp. 123–138.

Recanati, F. (2002). *Literal Meaning.* Cambridge: Cambridge University Press.

Renau, I. & Nazar, R. (2021). Verbario. URL: http://www.verbario.com. Last accessed: 1 June, 2021.

Romani, E. & Ježek, E. (2020). Tracing Metonymic Relations in T-PAS: An Annotation Exercise on a Corpus-based Resource for Italian. In *Proceedings of the 7th Italian Conference on Computational Linguistics*. Bologna, Italy.

Shutova, E., Kaplan, J., Teufel, S. & Korhonen, A. (2013). A computational model of logical metonymy. *ACM Transactions on Speech and Language Processing*, 11.