# Evaluating a rule based strategy to map IMAGACT and T-PAS

Andrea Amelio Ravelli Università di Firenze, Italy andreaamelio.ravelli @unifi.it Lorenzo Gregori Università di Firenze, Italy lorenzo.gregori @unifi.it

Anna Feltracco Fondazione Bruno Kessler Università di Pavia, Italy Università di Bergamo, Italy feltracco@fbk.eu

#### Abstract

**English.** This paper presents the analysis of a mapping between two resources, IMAGACT and T-PAS, made through a rule-based algorithm which converts argument structures in thematic roles. Results are good in terms of Recall, while Precision values are low: an analysis of the causes is proposed.

Italiano. Questo articolo presenta l'analisi di un mapping tra le risorse IMA-GACT e T-PAS, realizzato attraverso un algoritmo basato su regole che converte le strutture argomentali in ruoli tematici. I risultati sono buoni in termini di Recall, mentre sono bassi i valori di Precision per i quali viene proposta un'analisi.

## 1 Introduction

The automatic mapping of information between two resources is not a trivial task, but indeed joining information over specific data can benefit the involved resources. This paper describes the analysis of a mapping between two linguistic resources: IMAGACT and T-PAS. The motivation behind this mapping starts with the observation that both resources deal with Italian verbs disambiguation, are corpus-based and contain pieces of information that can be integrated with each other.

IMAGACT is a linguistic ontology of *actions*, that are grouped in *concepts* and related to different verb *Types*. For example, the action "John takes the cup from the shelf" belongs to the concept "*take an object*" and refers to Type 3 of the verb *to take*. Each Type is also associated to one or more thematic structures (e.g. [AGENT-verb-THEME-SOURCE]) and to videos via a set of captions.

T-PAS is a repository of *argument typed structures* for Italian verbs. Each verb is listed with its structures, which correspond to different senses of the verb. For each structure, the specification of the expected *semantic type* in every argument position (e.g. for the *subject*) is provided.

In this paper, we describe the results of a first attempt of mapping information between these resources. Specifically, for each of the 248 verbs analysed in both resources, we aim at matching the IMAGACT Types with the corresponding typed argument structures in T-PAS. We operate this mapping by applying a set of rules which convert the information from the argument structure into a thematic-role combination, and find all the Types that match this combination.

The linking between argument and thematic structures of a predicate is a debated complex task in linguistic theories (Baker, 1997; Pinker, 2009; Bowerman, 1990, among others). The predictability of thematic roles from argument structure (or viceversa) belongs to the *syntax-semantics interface*, and a study in this direction is out of the scope of this paper. Our experiment is focused on an empirical analysis of argument and thematic structures in Italian verbs and our aim is to evaluate whether, and to which extent, a rule-based system is able to produce thematic structures. We also intend to verify how these results can be exploited for a mapping purpose.

The paper is structured as follows: in Section 2 we present the resources; in Section 3 we describe the mapping procedure; in Section 4 we present and discuss the results of the mapping, tested on a gold standard; in Section 5 we provide direction for future work; in Section 6 we report our conclusions.

## 2 The Resources

In this section we describe IMAGACT and T-PAS. Table 1 shows the total and shared quantitative

	IMAGACT	T-PAS
Total Verbs	777	1,000
Total Types - t-pass	1,429	4,241
Shared Verbs	248	
Shared Types - t-pass	421	1,153

data of the two resources.

Table 1: Data of IMAGACT and T-PAS.

## 2.1 IMAGACT

IMAGACT<sup>1</sup> (Moneglia et al., 2014; Panunzi et al., 2014) is a visual ontology of action that provides a translation and disambiguation framework for action verbs. The resource contains a fine-grained categorization of action concepts, which are represented by one or more visual prototypes, in the form of recorded videos or 3D animations.

Action concepts are derived by a deep analysis of the most frequent action verbs in Italian and English spoken corpora: this ensures the ontology to cover the most relevant actions for our everyday activities. Given that no one-to-one correspondence can be established between an action verb and an action concept (Moneglia, 1993), each verb is divided in Types, which operate a segmentation of the predicate extension by identifying the prominent cores of the verb meaning. Verb Types are connected to action concepts and they are the linkage point between lexical and action levels (Moneglia et al., 2012a). Types in IMA-GACT are inter-connected through semantic relations and gather the sentences retrieved in the spoken corpora, which have been classified and linguistically annotated with thematic roles and aktionsart<sup>2</sup>.

The resource is growing continuously: by now, it consists of a total of 1010 action concepts, each one with a visual representation (i.e. a scene), and 21 covered languages (9 fully-mapped, 13 underway), with an average of 730 action verbs per language.

## 2.2 T-PAS

T-PAS<sup>3</sup>, Typed Predicate Argument Structures (Jezek et al., 2014), is a repository of verb patterns acquired from corpora by manual clustering distributional information about Italian verbs. For every

<sup>3</sup>http://tpas.fbk.eu/

typed structure (henceforth *t-pas*), the specification of the expected semantic type (ST) for each argument slot is provided. T-PAS accounts for the following argument positions: *subject, object, indirect object, complement, adverbial* and *clausal*. A description of the sense, in the form of an *implicature*, is also linked to the *t-pas*.

Example 1 reports the *t-pas#2* of the verb *ab-battere*: the STs [[Human]] and [[Event]] are specified for the subject position (as alternatives) and [[Building]] for the object position.

(1) [[Human | Event]-subj] abbattere [[Building]-obj] implicature:[[Human | Event]] distrugge, butta giù [[Building]] example: "Il muratore abbatte la parete." (Eng. "The bricklayer knocks the wall.")

The STs aim at generalizing over the set of lexical items observed in a certain position for a particular sense of the verb. For instance, in Example 1, the ST [[Building]] generalizes over the lexical item *parete* (Eng. *wall*). STs are drawn from a list of about 230 types<sup>4</sup> and are also organized in a hierarchy, in which the elements are linked by a "IS-A" relation (Jezek et al., 2016). Table 2 presents a section of the hierarchy in which it is shown that [[Plane]] IS-A [[Vehicle]], [[Vehicle]] IS-A [[Machine]] and so on.<sup>5</sup> If no generalization is possible, the set of lexical items found in the argument position is listed.

Table 2: Section of the STs hierarchy.

Each *t-pas* corresponds to a distinct sense of the verb and is identified and defined by analysing instances of the verb in a corpus, following the lexicographic procedure called Corpus Pattern Analysis (Hanks, 2004; Hanks and Pustejovsky, 2005).<sup>6</sup> The corpus instances are then associated to the corresponding *t-pas*.

<sup>&</sup>lt;sup>1</sup>http://www.imagact.it/

<sup>&</sup>lt;sup>2</sup>See Moneglia et al. (2012b) for details on annotated data and ontology building process.

<sup>&</sup>lt;sup>4</sup>For details on the list creation see (Jezek et al., 2014).

<sup>&</sup>lt;sup>5</sup>The same list has been used for the English resource PDEV (Hanks and Pustejovsky, 2005), http://pdev.org.uk. The hierarchy can be found in http://pdev.org.uk/#onto.

<sup>&</sup>lt;sup>6</sup>According to the CPA procedure, after analysing a random sample of 250 concordances of the verb in the corpus, each *t-pas* is defined by recognizing its relevant structure and identifying the STs for each argument slots.



Figure 1: An example of the mapping between IMAGACT and T-PAS for the verb macinare.

T-PAS currently contains 1000 verbs. The reference corpus is a reduced version of ItWAC (Baroni and Kilgarriff, 2006).

## 3 The Mapping

We aim at finding the best semantic match between a verb Type in IMAGACT and the *t-pass* of the same verb in T-PAS, the two referring to the same action concept. Notice that it is possible that a Type in IMAGACT is mapped to more than one *t-pas* due, for instance, to different possible verb alternations that can occur inside the same Type. Figure 1 shows an example of this mapping, in which there is a match between Type 1 and tpas#1 of the verb *macinare*.

The mapping is done as follows. By observing a sample of verbs in the resources, we first defined a set of simple rules to convert the t-pas in a thematic structure. Considering the ST in the argument positions of the *t-pas* (e.g. [Human]-subj, [Food]-obj]), the rules aim at creating a thematic structure for the t-pas of the kind AG-v-TH (dotted arrow in Figure 1). Then, we used an algorithm which applies these rules to all the *t*-pass of a verb, and map the derived thematic structure (derivedts) to the thematic structures (ts) of the Types in IMAGACT (horizontal arrow in Figure 1). The system thus compares all the ts in IMAGACT with all the derived-ts in T-PAS for the same verb, and retrieves the matches.<sup>7</sup> In Figure 1, the *t-pas#1* for the verb macinare have been transformed in the structure AG-v-TH and then mapped to the ts of the Type.

The mapping between IMAGACT and T-PAS is made for the 248 verbs common to the two resources.

**Datasets** The rules for the conversion of a *t*-*pas* in a *derived-ts* have been manually created by observing a sample of 15 verbs shared by the two resources (devset). We evaluated the mapping against a gold standard manually created by pairing the Types of other 14 verbs with the corresponding *t*-*pass*. We extracted the 29 verbs from the 248 shared by the two resources. The selection was made preserving the variability of the verbs in the two resources, in terms of their number of Types or *t*-*pas*. For instance, *prendere* (*to take*) is associated with 17 *t*-*pass* in T-PAS and 18 Types in IMAGACT; on the contrary *bussare* (*to knock*) has only 2 *t*-*pass* and 1 Type.

**Conversion rules** Table 3 synthesizes the rules we adopted. The rules consider both the ST in the argument slot and the argument slot itself, and are meant to associate a ST in an argument slot to a thematic role. For example, line 7 of Table 3 has to be interpreted as follows: if for the subject position of the *t-pas* the ST [[Animate]] (or a IS-A [[Animate]], according to the hierarchy of ST) is expected, then the AGENT role is selected (line 8). The rules also consider if the verb is in reflexive form (line 13). Moreover, if the *t-pas* registers the ST [[Abstract Entity]] (or a ST that IS-A [[Abstract Entity]]) as unique ST for any argument position (i.e. it is the only ST expected for the position), the *t-pas* was excluded from the mapping, as IMAGACT only accounts for physical actions which do not involve abstract entities.

#### 4 Results and discussion

In order to calculate Precision (P) and Recall (R) of the algorithm, we considered that DESTI-NATION (DE), SOURCE (SO) and LOCATION (LO) roles can not always be discriminated (for example, *room* is a DE in "John puts a table in the room", a SO in "John takes the table from

<sup>&</sup>lt;sup>7</sup>Notice that the mapping is considering just this information of the resources and does not consider e.g. captions in IMAGACT or examples in T-PAS.



Table 3: Rules for mapping.

the room", a LO in "John walks in the room"). The same happens for AGENT (AG) and ACTOR (AC): a human can be an agent ("John sweeps the room") or an actor ("John bumps his head"). These limits can not be exceeded by an improvement of the rule definitions, because they are strictly dependent on the verb semantics. When calculating P and R, we grouped these derived structures together.

Precision (P)	Recall (R)	F-measure (F1)
0.283	0.792	0.418

Table 4: Precision, Recall, F1 of the mapping.

We observe good values for R, while the P is very low (Table 4). A deeper analysis shows that in 34.61% of the cases, we have a full match with the gold standard and in 38.46% the results from the mapping include the ones expected by the gold standard. This means that in many cases the system is able to retrieve the correct matches.

Figure 2 shows the distribution of the main thematic structures in the Types of the whole IMA-GACT ontology (in orange), in the devset (in red), compared with the *derived-ts* from T-PAS (in green). We verified *a posteriori* that the distribution of *ts*s in the devset is strictly comparable with the one in the whole ontology, meaning that the devset is also well-balanced in terms of the thematic structures coverage (see orange and red bars in Figure 2).

By using the transformational rules we were able to recreate all the structures that are used in IMAGACT; however, there are some discrep-



Figure 2: Distribution of the thematic structures.

ancies in the production of AG-v-TH, TH-v (too high) and AG-v-TH-[DE|LO|SO] (too low) (see Figure 2).

The critical issue is represented by the AG-v-TH structure: this is the most frequent one among the IMAGACT Types and in our test set (112 over 166 Types). For example, the following sentences belong to 4 different Types of the verb stringere, but have the same ts AG-v-TH: "Marco stringe la mano a Luca"; "Marco stringe le gambe"; "Marco stringe i pugni"; "Marco stringe la vite". This happens also for the t-pas of stringere: 3 over the 5 derived-ts are AG-v-TH, so the system produces 12 combinations over 3 attested in the gold standard. The high frequency of this structure strongly influences the final P and R results. Moreover, the ts AG-v-TH is not distinctive of Types intra-verbs: by taking all the verbs with more than one Type, and for which AG-v-TH is a possible ts, we measured that in only 38,22% of them this ts is present in only one Type; in the other verbs (61.78%) the AG-v-TH structure appears in more than one Type.

#### 5 Future work

Given the result in terms of Precision we presented in the previous section, we are considering to adopt other strategies that can be useful for the mapping of IMAGACT and T-PAS.

For instance, it would be possible to exploit the examples from the corpus associated with each *t*-*pas* in T-PAS. In this sense, we hypothesize the processing of these examples through BabelFy (Moro et al., 2014), an online system for word sense disambiguation, based on the BabelNet semantic network (Navigli and Ponzetto, 2012). BabelNet is already linked to IMAGACT (via the scenes). We can use BabelFy in order to perform the disambiguation of a verb in the sentences associated to each *t*-*pas*. In this way we can ob-

tain a link between the verb under examination and the corresponding BabelNet synset (i.e., a Babel-Synset). The application of this method to every example will result in a ranking of the most frequent BabelSynsets for the group of sentences of each *t-pas*. Combining this output ranking with the BabelNet-IMAGACT linking (Gregori et al., 2016), we will obtain the set of IMAGACT Types that most likely match with each *t-pas*.

On the other way round, IMAGACT captions could also be mapped into the corresponding tpass, by using the output of the algorithm developed in (Feltracco et al., 2016): given a sentence of a t-pas, the algorithm identifies the lexical item(s) that are generalized by the ST for each argument position of every *t-pas* (e.g. assigning the ST [[Building]] to "parete" in the sentence "II muratore abbatte la parete" for the *t-pas* [[Human ] Event]] abbattere [[Building]]). A measure of semantic similarity between the lexical items of an IMAGACT caption and the set of items associated to the same verb in T-PAS, would provide an approximation of which are *t-pass* that most likely match the given caption. The application of this method to every caption of an IMAGACT Type will help us in the goal of mapping T-PAS with IMAGACT.

This method added to our rule-based strategy can be particularly useful to solve the ambiguity related to the thematic pattern AG-v-TH, for which the use of lexical information would reduce the number of possible matches.

## 6 Conclusions

In this paper we presented a first attempt of mapping IMAGACT and T-PAS by using a rule-based algorithm for the automatic conversion of T-PAS semantic types into thematic structures. We took advantage of the strong discriminative power of semantic types in their argument position to reduce the possible set of allowed thematic structures. This approach has an intrinsic limit: thematic roles are determined by verb semantics and their difference is not always reflected in the related semantic type. We also found out that the *ts* AG-v-TH represents the most critical issue, being the most frequent structure, and appearing in more than one Type of the same verb.

The results report a good recall and a low precision, confirming that our algorithm is not able to produce an actual mapping between the two resources, but it provides a reliable set of mapping candidates: we believe that it can be fruitfully exploited for a first step of a mapping process, in order to filter a lot of unwanted matching possibilities. We are confident that by exploiting additional linguistic information from the two resources (e.g. captions and occurrences in IMAGACT, lexical information and examples in T-PAS), the precision of this mapping will improve sensibly.

#### References

- Mark C Baker. 1997. Thematic roles and syntactic structure. In *Elements of grammar*, pages 73–137. Springer.
- Marco Baroni and Adam Kilgarriff. 2006. Large linguistically-processed web corpora for multiple languages. In *Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics: Posters & Demonstrations*, pages 87–90. Association for Computational Linguistics.
- Melissa Bowerman. 1990. Mapping thematic roles onto syntactic functions: are children helped by innate linking rules? *Linguistics*, 28(6):1253–1290.
- Anna Feltracco, Lorenzo Gatti, Simone Magnolini, Bernardo Magnini, and Elisabetta Jezek. 2016. Using WordNet to Build Lexical Sets for Italian Verbs. In Proceedings of the Eighth Global WordNet Conference (GWC '16), Bucharest, Romania, January.
- Lorenzo Gregori, Alessandro Panunzi, and Andrea Amelio Ravelli. 2016. Linking IMAGACT ontology to BabelNet through action videos. Proceedings of Third Italian Conference on Computational Linguistics (CLiC-IT 2016), pages 162–167.
- Patrick Hanks and James Pustejovsky. 2005. A pattern dictionary for natural language processing. *Revue française de linguistique appliquée*, 10(2):63–82.
- Patrick Hanks. 2004. Corpus pattern analysis. In *Proceedings of the Eleventh EURALEX International Congress, Lorient, France, Universite de Bretagne-Sud.*
- Elisabetta Jezek, Bernardo Magnini, Anna Feltracco, Alessia Bianchini, and Octavian Popescu. 2014. T-PAS: a resource of corpus-derived types predicateargument structures for linguistic analysis and semantic processing. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, May. European Language Resources Association (ELRA).
- Elisabetta Jezek, Anna Feltracco, Lorenzo Gatti, Simone Magnolini, and Bernardo Magnini. 2016. Mapping Semantic Types onto WordNet Synset. In

Proceedings of the Twelfth Joint ACL - ISO Workshop on Interoperable Semantic Annotation (Isa '12), Portorose, Slovenia, May.

- Massimo Moneglia, Gloria Gagliardi, Lorenzo Gregori, Alessandro Panunzi, Samuele Paladini, and Andrew Williams. 2012a. La variazione dei verbi generali nei corpora di parlato spontaneo. L'ontologia IMAGACT. In Proceedings of the VIIth GSCP International Conference: Speech and Corpora, pages 406–411.
- Massimo Moneglia, Gloria Gagliardi, Alessandro Panunzi, Francesca Frontini, Irene Russo, and Monica Monachini. 2012b. Imagact: deriving an action ontology from spoken corpora. In *Eighth Joint ACL-ISO Workshop on Interoperable Semantic Annotation (isa-8)*, pages 42–47.
- Massimo Moneglia, Susan Brown, Francesca Frontini, Gloria Gagliardi, Fahad Khan, Monica Monachini, and Alessandro Panunzi. 2014. The IMAGACT Visual Ontology. An Extendable Multilingual Infrastructure for the Representation of Lexical Encoding of Action. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, May. European Language Resources Association (ELRA).
- Massimo Moneglia. 1993. Prototypical vs. nonprototypical predicates: ways of understanding and the semantic partition of lexical meaning. In *International conference*" *Linguistics at the end of the century*" *Moscow State University February*.
- Andrea Moro, Alessandro Raganato, and Roberto Navigli. 2014. Entity Linking meets Word Sense Disambiguation: a Unified Approach. *Transactions* of the Association for Computational Linguistics (TACL), 2:231–244.
- Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217– 250.
- Alessandro Panunzi, Irene De Felice, Lorenzo Gregori, Stefano Jacoviello, Monica Monachini, Massimo Moneglia, Valeria Quochi, and Irene Russo. 2014. Translating Action Verbs using a Dictionary of Images: the IMAGACT Ontology. In XVI EU-RALEX International Congress: The User in Focus, pages 1163–1170, Bolzano / Bozen, 7/2014. EU-RALEX 2014, EURALEX 2014.
- Steven Pinker. 2009. Language learnability and language development, with new commentary by the author, volume 7. Harvard University Press.