

Defining Action Types: on the roles of Thematic Structures and Argument Alternations

Massimo Moneglia
Università di Firenze

Alessandro Panunzi
Università di Firenze

Rossella Varvara
Università di Firenze

massimo.moneglia@unifi.it alessandro.panunzi@unifi.it rossella.varvara@unifi.it

Abstract

English. The paper focuses on the relation between Action Types (ontological objects that identifies the referential potential of a verb) and the Thematic Structure and Alternations of verbs. The empirical analysis shows that these linguistic features are not properties of the verb itself, but vary in relation to its referential variation. Given this evidence, we argue that Thematic Structure and Argument Alternation can help in the identification of the different types of action to which a verb refers, providing evidences to define the granularity of action concepts in the development of an ontology of actions.

1 Introduction

Action verbs are among the most frequent words in ordinary communication, and their correct processing constitutes an underpinning element for a wide series of human-machine interaction tasks. The formalization of action verb meanings has often been linked to propositional representations within decompositional approaches (Dowty, 1979; Rappaport Hovav and Levin, 2012), in which the semantic core of the verb remains a non-analyzed lexical root. Other traditional components used in the representation and annotation of the meaning of action verbs are: the temporal and aspectual properties of verbs (Vendler, 1957; Pustejovsky, 1991); the thematic roles of participants (Fillmore, 1967; Gruber, 1965); the force dynamics and causal relations implied (Talmy, 1988; Croft, 2012; Gärdenfors, 2014). Nevertheless, even if these semantic components are usually assumed to reflect the general structure

of action conceptualization, the linguistic and the cognitive levels of categorization are not equivalent and should not be confused (Croft, 2012; Moneglia, 2014). As a matter of fact, the lexical category instantiated by an action verb can refer to more than one cognitive entity.

For instance, the verb *to push* can refer to actions in which the force causes the movement of the object in space (e.g. in a sentence like *John pushes the basket under the table*), as well as to actions in which the object does not move (e.g. *John pushes the fabric into a ball*). This differential property is more than enough to cognitively distinguish these events in different action concepts. As a consequence, the need for a cognitive level of action categorization which is independent from the lexical one becomes clear.

In this paper, we investigate the role of one type of linguistic information, specifically Thematic Structure and Argument Alternations, in the definition of action types, i.e. types of action concepts that gather together single datapoint in the IMAGACT ontology of actions. We point out that Thematic Structure is not a property of the verb itself, since different structures may be present for the same verb. Our aim is to show how these features are linked to action types and how this correlation can be useful for the definition and the induction of Action Types¹.

In section 2, we show the innovative methodology assumed by the IMAGACT Ontology of Action for representing the meaning of action verbs, focusing on their referential properties rather than on their intensional definition. In sections 3 and 4, we will see through a case study that the induction of the referential variation of verbs can take advantage of linguistic features. Thematic Struc-

Copyright ©2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹Similarly, previous work in Word Sense Disambiguation (Dang and Palmer, 2005; Roberts and Kordoni, 2012) have shown that thematic information can improve verb sense disambiguation.

tures and their Alternations can have an impact in the definition and characterization of the objects in an ontology of actions.

2 The IMAGACT ontology

In the IMAGACT multilingual Ontology of Actions² (Moneglia et al., 2012b; Panunzi et al., 2014) action concepts are not defined through a propositional and truth conditional perspective, but they are rather identified and visually represented through scenes. Each scene is conceived as a prototypical instance (Rosch, 1983) of an action concept and constitutes the basic entity of reference of the action ontology. Scenes have been derived from a complex annotation procedure (Moneglia et al., 2012a) of the occurrences of action verbs³ in two large spoken resources of English and Italian. After this bootstrapping phase, the ontology has been extended to many other languages exploiting competence judgments by native speakers (Brown et al., 2014; Pan et al., 2018; Moneglia et al., 2018b). The whole IMAGACT database is currently comprised of 1,010 scenes linked to more than 8,700 verbs in 13 languages⁴. As a result, action concepts have been represented by language independent scenes, each one linked to the series of verbs that can be used to refer to it. The scenes are described by linguistic captions (i.e. short sentences) that have as predicates each of those verbs. Simultaneously, each verb is connected to a set of scenes in the ontology, representing in this way its referential variation.

The scenes linked to a verb have been then grouped in broader categories called Action Types⁵ (hereafter also *ATs* or *Types*). ATs are defined as higher level concepts which fall in the semantic variation of a verb, useful to represent its referential potential in a more compact way, reducing an excessive granularity in the representation of meaning⁶. ATs have been created exploiting similarity judgments among scenes and considering Local Equivalent verbs, i.e. all the verbs

that could co-extend to the same scene (Moneglia et al., 2018a). An additional validation, in which raters were asked to assign scenes to ATs, was conducted with an overall agreement of 0.8 (Gagliardi, 2014). Lastly, during the ontology’s development, Thematic Structure, Alternations and Aktionsart were manually annotated for the linguistic captions of each scene. These latter annotations will be the starting point of the present study, in which we analyze the correlation between ontological entities (ATs) and linguistic features, specifically Thematic Structure and Alternations.

3 Thematic Structure and Action Types

Thematic structure and syntactic frame information of verbs are usually provided by most lexical resources, such as VerbNet (Kipper-Schuler, 2005), FrameNet (Fillmore et al., 2004) and PropBank (Palmer et al., 2005). In these resources, the different entries of a verb are associated to their possible thematic structures. They include manually annotated data and have been useful for the development of statistical approaches for Semantic Role Labeling (Gildea and Jurafsky, 2002) and for various NLP applications (e.g. information extraction (Surdeanu et al., 2003), summarization (Melli et al., 2006), and machine translation (Boas, 2002)).

In this section, we show that Thematic Structure (TS) is not a property of the verb and we will verify: 1) to what extent it can be considered a property of the action types in the variation of a verb; 2) to what extent it can provide a differential feature for the identification of ATs. We consider as TS the minimal thematic structure⁷ which is necessary to interpret a verb as an instance of a specific AT.

There are cases in which the TS is the same all through the verb variation. Frequently, one specific thematic structure is associated to activity verbs that show almost no variation in their meaning. This is the case of the verb *to drink*, who has only one AT. The verb *to close*, on the contrary, shows a significant variation in the IMAGACT ontology (7 ATs, four of them represented in table 1), but all types present the same TS (AG-V-TH). In these cases, thematic structure cannot play any role in the definition of different types, which are

²Freely accessible at <http://www.imagact.it/>

³Only in their basic, physical meaning, so excluding all metaphorical and phraseological uses.

⁴Besides English and Italian, the list of fully mapped language comprehends: Arab, Chinese, Danish, German, Hindi, Japanese, Polish, Portuguese, Serbian, Spanish, Greek.

⁵See, for instance, Table 1 which represents the main variation of the action verb *to close*.

⁶As a matter of fact, some verbs in IMAGACT can be linked to several dozen scenes, and the most general ones, like *to take* and *to put*, are linked to about 100 scenes.

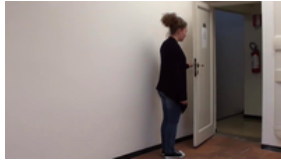
⁷The set of roles used in IMAGACT is based mainly on the set used in VerbNet.



Mary closes her hand



Mary closes the umbrella



Mary closes the door



Mary closes the lock

Table 1: Variation of *to close*

identified on the basis of ontological features of the theme (e.g. a body part vs an artifact) or by the kind of result produced (spatial consequences vs functional consequences), and even by the set of equivalent verbs which provide a differential property of each ATs (*to shut* vs. *to lock* vs. *to close up* vs. *to clench*).

Verbs like *to close* shows that TS is not a necessary differential of ATs, but, as the next examples will point out, it can help to select among the interpretations of a general verb. This is the case with verbs like *to press* and *to push* which record different TSs possibilities across their variation. Let's consider the verb *to press*. In the IMA-GACT ontology it shows 10 different ATs. We can observe groups of Types that share the same TS. Types *a* (table 2) and *b* (table 3) present Agent-Verb-Theme-Destination structure. In both cases, the destination is necessary to represent the event type, which cannot be identified otherwise. In type *a*, the Agent compacts the Theme into a block, changing its shape but not its volume. In type *b* the Agent squeezes the Theme, reducing its volume.



John presses the scraps into a block
AG-V-TH-DEST
to push, to compact

Table 2: *To press*, type *a*

Types *c*, *d* and *e* (tab.4, 5, 6) differ from types *a* and *b* since Destination is not necessary and AG-V-TH is sufficient to identify the action.

Despite the common Thematic Structure, they clearly identify different actions for cognitive reasons. In type *c* the Theme is a humans body part,



Mary presses the fabric into a ball
AG-V-TH-DEST
to push, to squeeze, to compress

Table 3: *To press*, type *b*



The doctor presses the shoulder
AG-V-TH
to push, to poke

Table 4: *To press*, type *c*



John presses the button
AG-V-TH
to push

Table 5: *To press*, type *d*



John presses the pedal
AG-V-TH
to push

Table 6: *To press*, type *e*

and the concept implies a compression as an inter-subjective activity, whereas in type *d* the Theme is an object and the compression implies a functional correlation. In type *e* the Theme undergoes a continuous scalar pressure, not limited to a single impulse.

Although these TS commonalities among types show that TS is not necessarily predictive of a single type, TS helps in the distinction of action types. For example, TS restricts the range of possible interpretation of a general verb like *to press* in the case of type *a* and type *b* (table 2 and 3) versus type *c*, *d* and *e* (tables 4, 5 and 6). The distinction between these groups of types (which is independently motivated) is mirrored by the different TSs.

TS may constitute an important feature for the definition of granularity of action types in the verb variation. Type *c* (*the doctor presses the shoulder*, tab.4) and type *f* (*the thief presses the gun into Marys back*, tab.7) are distinguished in reason of their TS: they are similar actions from a cognitive point of view and they can be paraphrased both with *to push*, but the TS of the verb in the

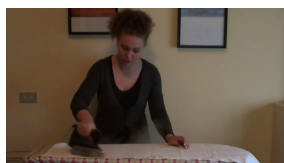
two events is different.



The thief presses the gun into Mary's back
AG-V-TH-DEST
to push

Table 7: *To press*, type *f*

Two cases in the variation of *to press* are characterized by a specific TS: type *g* (AG-V-TH-INSTR) and type *h* (AG-V-TH-RESULT). Type *g* (tab. 8) necessarily requires the instrument in its minimal structure, contrary to all other types; type *h* (tab. 9) requires the expression of the result of the action. TS is predictive of the Action Type in those cases.



The tailor presses the cloth with the iron
AG-V-TH-INSTR
to push

Table 8: *To press*, type *g*



John presses the can flat
AG-V-TH-RESULT
to push

Table 9: *To press*, type *h*

Considering the variation of a verb like *to press*⁸, we can conclude that TS is not peculiar of the verb but is related to its types. Given the cases in which one TS is shared by multiple types, it is clear that types distinction is not a function of the thematic variation (which is actually related to the intersection of multiple features). However, TS has a role in type prediction, since it helps identifying the features of a type.

4 The role of Argument Alternation

Argument Alternations (AAs) are one of those properties of the verb that have received great attention in a large body of literature after Levin (1993). As we will see, also AAs are not properties of the whole verb, but rather characterizes the verb in its types. Considering the verb *to press*, it

⁸Further similar examples have been extracted from the IMAGACT ontology; however, for space limitations, we refer only to the cases already discussed.

has been previously stated that it shows the *conative alternation*, i.e. “a transitivity alternation in which the objects of the verb in the transitive variant turns up in the intransitive conative variant as the object of the preposition in a prepositional phrase headed by the preposition *at* (sometimes *on* with certain verb of ingesting and the push/pull verbs)” (Levin, 1993, p.42). However, only four types of *press* allow for the *conative alternation*, as illustrated in the examples below:

- Type *c*: The doctor presses the shoulder → The doctor presses on the shoulder
- Type *d*: John presses the button → John presses on the button
- Type *e*: John presses the pedal → John presses on the pedal

Other types result in agrammatical sentences when the *conative alternation* is applied:

- Type *a*: *John presses at/on the scraps into a block
- Type *g*: *The tailor presses at/on the cloth with the iron

Considering now *to push*, a verb that shares many interpretations with *to press*, only some types of *to push* (types *a*, *b*, *c*, *d* but not *e*, *f* and *g*) allow this alternation:

- Type *a*: John pushes the button → John pushes on the button
- Type *b*: John pushes the shoulder → John pushes on the shoulder
- Type *c*: John pushes the lever → John pushes on the lever
- Type *d*: John pushes the pedal → John pushes on the pedal
- Type *e*: Mary pushes the chair → *Mary pushes on the chair
- Type *f*: Mary pushes the toothpaste → *Mary pushes on the toothpaste
- Type *g*: Mary pushes the fabric → *Mary pushes on the fabric

In addition to the conative alternations, other two alternations can be seen in the variation of the verbs considered: the *resultative construction* and the *theme-instrument alternation*. The *resultative construction* presents a phrase that describes the state achieved by the referent of a noun phrase as a result of the action. As noted already by Levin (1993, p. 100), it cannot be predicated in case of oblique:

- (1) a. The silversmith pounded the metal flat.
- b. *The silversmith pounded on the metal flat.

This alternation is found only in type *h* for *to press*:

- John presses the can → John presses the can flat

Lastly, we find an alternation between the Theme and the Instrument, not listed in Levin (1993). In this case, the Instrument from sentence 2b (which can be seen as the result of a conative alternation) becomes the Theme in sentence 2c.

- (2) a. The doctor pushes the shoulder with his hand
- b. The doctor pushes on the shoulder with his hand
- c. The doctor pushes his hand on the shoulder

This alternation can be considered as a particular case of *locative alternation*. In terms of Levin (1993), the noun *shoulder* would represent the location argument, whereas *hand* would be considered the *locatum*. Also in this case, the *theme-instrument alternation* does not apply to all types of the variation of *to press*, but rather characterizes specific types.

- Type *g*: the tailor presses the clothes with the iron → The tailor presses the iron on the clothes
- Type *c*: the doctor presses the shoulder → the doctor presses the shoulder with the hand → the doctor presses the hand on the shoulder
- Type *d*: John presses the button → John presses the button with the hand → John presses the hand on the button

- Type *f*: the thief presses the gun into Marys back → the thief presses Marys back with the gun⁹

For the verb *to push*, only types *i* and *d* allow it:

- Type *i*: The thief pushes the gun into Marys back → The thief pushes Marys back with the gun
- Type *d*: John pushes the pedal → John pushes the foot on the pedal

As a whole, if considered together, TS and AA can reduce the underdetermination of types. In other words, when two types share the same TS, they can be predicted from a difference in their argument alternations. This is the case, for example, for types *a* (table 2) and *f* (table 7) of *to press*, which share the TS AG-V-TH-DEST, but differ with respect to the theme-instrument alternation: only type *f* allows it, not type *a*.

In the variation of *to push*, types *e* and *a* share the same TS (AG-V-TH) but type *e* does not allow the conative alternation (*≠Mary pushes on the chair*), contrary to types *a* (*John pushes on the button*).

5 Conclusion

In this paper we have investigated the relation between Thematic Structure and Action Types. The empirical analysis reveals that Thematic Structure and Argument Alternations are not properties of the whole verb, but rather of the verb in its Types. We have provided evidence about the saliency of both Thematic Structure and Argument Alternations in the identification of Action Types. Although TS and AA do not determine the variation of a verb across different ATs, these linguistic features can, indeed, reveal characterizing features of a Type, helping us in the disambiguation of concepts and in the recognition of the necessary level of granularity in building our ontologies.

References

Hans Christian Boas. 2002. Bilingual framenet dictionaries for machine translation. In *Proceedings of LREC*.

⁹Other types do not allow the theme-instrument alternation: *John presses the hand on the scraps (type *a*); *Mary presses the hand on the clothes (type *i*). For completeness, we report some examples of *to push* for which this alternation does not hold: *Mary pushes on the chair with her hand (type *e*); *Mary pushes the hand on the box (type *h*).

- Susan Windisch Brown, Gloria Gagliardi, and Massimo Moneglia. 2014. Imagact4all. mapping spanish varieties onto a corpus-based ontology of action. *CHIMERA: Journal of Romance Corpora and Linguistic Studies*, (1):91–135.
- William Croft. 2012. *Verbs: Aspect and causal structure*. OUP Oxford.
- Hoa Trang Dang and Martha Palmer. 2005. The role of semantic roles in disambiguating verb senses. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 42–49. Association for Computational Linguistics.
- David Dowty. 1979. *Word Meaning and Montague Grammar*. Reidel Publishing Co, Dordrecht.
- Charles J Fillmore, Josef Ruppenhofer, and Collin F Baker. 2004. Framenet and representing the link between semantic and syntactic relations. *Frontiers in linguistics*, 1:19–59.
- Charles J. Fillmore. 1967. The case for case. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*, pages 1–89. Holt, Rinehart and Winston, New York.
- Gloria Gagliardi. 2014. *Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnosi del Mild Cognitive Impairment*. Ph.D. thesis, University of Florence.
- Peter Gärdenfors. 2014. *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press, Cambridge (MA).
- Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational linguistics*, 28(3):245–288.
- Jeffrey Gruber. 1965. *Studies in Lexical Relations*. Ph.D. thesis, M.I.T.
- Karin Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, University of Pennsylvania.
- Beth Levin. 1993. *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Gabor Melli, Yang Wang, Yudong Liu, Mehdi M Kashani, Zhongmin Shi, Baohua Gu, Anoop Sarkar, and Fred Popowich. 2006. Description of squash, the sfu question answering summary handler for the duc-2005 summarization task. *Proceedings of the HLT/EMNLP Document Understanding Workshop (DUC)*.
- Massimo Moneglia, Gloria Gagliardi, Alessandro Panunzi, Francesca Frontini, Irene Russo, and Monica Monachini. 2012a. Imagact: Deriving an action ontology from spoken corpora. In *Proceedings of the Eight Joint ACL - ISO Workshop on Interoperable Semantic Annotation (ISA-8)*. Pisa, October 3-5, 2012, pages 42–47.
- Massimo Moneglia, Monica Monachini, Omar Calabrese, Alessandro Panunzi, Francesca Frontini, Gloria Gagliardi, and Irene Russo. 2012b. The imagact cross-linguistic ontology of action. a new infrastructure for natural language disambiguation. In Nicoletta Calzolari, editor, *Proceedings of the Eight International Conference on Language Resources and Evaluation*, pages 948–955. European Language Resources Association (ELRA).
- Massimo Moneglia, Alessandro Panunzi, and Lorenzo Gregori. 2018a. Action identification and local equivalence of action verbs: the annotation framework of the imagact ontology. In James Pustejovsky and Ielka van der Sluis, editors, *Proceedings of the LREC 2018 Workshop AREA Annotation, Recognition and Evaluation of Actions*, pages 23–30. European Language Resources Association (ELRA).
- Massimo Moneglia, Alessandro Panunzi, and Lorenzo Gregori. 2018b. Taking events in hindi. a case study from the annotation of indian languages in imagact. In *Proceedings of the LREC 2018 Workshop WILDRE4 4th Workshop on Indian Language Data: Resources and Evaluation*, pages 46–51. LREC.
- Massimo Moneglia. 2014. Natural language ontology of action: A gap with huge consequences for natural language understanding and machine translation. In Z. Vetulani and J. Mariani, editors, *Human Language Technology. Challenges for Computer Science and Linguistics.*, pages 370–395. Springer, Berlin/Heidelberg.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.
- Yi Pan, Massimo Moneglia, Alessandro Panunzi, and Lorenzo Gregori. 2018. Imagact4all. una ontologia per immagini dell'azione per l'apprendimento del lessico verbale di base delle lingue seconde. In Anna De Meo and Margaret Rasulo, editors, *Usare le lingue seconde*, pages 120–150. Officinaventuno.
- Alessandro Panunzi, Irene De Felice, Lorenzo Gregori, Stefano Jacoviello, Monica Monachini, Massimo Moneglia, and Valeria Quochi. 2014. Translating action verbs using a dictionary of images: the imagact ontology. In *Proceedings of the XVI EU-RALEX International Congress: The User in Focus. Bolzano: EURAC research*, pages 1163–1170.
- James Pustejovsky. 1991. The syntax of event structure. *Cognition*, 41:47–81.
- Malka Rappaport Hovav and Beth Levin. 2012. Building verb meanings. In Miriam Butt and Wilhelm Geuder, editors, *The projection of arguments: Lexical and compositional factors*, pages 97–134. CSLI Publications, Stanford, CA.
- Will Roberts and Valia Kordoni. 2012. Using verb subcategorization for word sense disambiguation. In *LREC*, pages 829–832.

Eleanor Rosch. 1983. Prototype classification and logical classification: The two systems. *New trends in conceptual representation: Challenges to Piaget's theory*, pages 73–86.

Mihai Surdeanu, Sanda Harabagiu, John Williams, and Paul Aarseth. 2003. Using predicate-argument structures for information extraction. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*.

Leonard Talmy. 1988. Force dynamics in language and cognition. *Cognitive science*, 12(1):49–100.

Zeno Vendler. 1957. Verbs and times. *The philosophical review*, 56:97–121.