

From Visual Prototypes of Action to Metaphors Extending the IMAGACT Ontology of Action to Secondary Meanings

Susan Windisch Brown

University of Florence
Piazza Savonarola, 1, Florence, Italy
E-mail: susanwbrown@att.net

Abstract

This paper describes an infrastructure that has been designed to deal with corpus-based variations that do not fall within the primary, physical variation of action verbs. We have first established three main categories of secondary variation--*metaphor*, *metonymy* and *idiom*--and criteria for creating types within these categories for each verb. The criteria rely heavily on the images that compose the IMAGACT ontology of action and on widely accepted processes of meaning extension in linguistics. Although figurative language is known for its amorphous, subjective nature, we have endeavoured to create a standard, justifiable process for determining figurative language types for individual verbs. We specifically highlight the benefits that IMAGACT's representation of the primary meanings through videos brings to the understanding and annotation of secondary meanings.

Keywords: semantic annotation, metaphor, metonymy

1. Introduction

IMAGACT is a cross-linguistic ontology of action concepts that are represented with prototypic 3D animations or brief films. This format makes use of the universal language of images to identify action types, avoiding the under-determinacy of semantic definitions. This ontology has been induced from the references to physical actions found in English and Italian spoken corpora (Moneglia et al. 2012) and gives a picture of the variety of activities that are prominent in our everyday life, specifying the language used to express each one in ordinary communication.

IMAGACT uses prototypic scenes to represent the range of variations that natural language verbs can record in a language and maps different languages onto the same ontology of visually represented concepts. Each verb can express one or more concepts, while each concept can refer to one or more verbs. (Moneglia in press).

For example, the verb *to cross* ranges over four main action types (Figure 1), identified in corpus occurrences, some of which can be equivalently identified by other verbs (*pass*, *climb*). The specific way of categorizing actions by the verb *to cross* does not find direct correspondence in other languages. For instance, in Italian only type 1 and 3 can be in the extension of the direct translation (*attraversare*) while 2 and 4 respectively require other Italian verbs (*incrociare*, *superare*).

The IMAGACT ontology has been developed through annotation of English and Italian spoken corpora, in which reference to actions is frequent. Working in their native languages, linguists identified the variation of action-oriented lexicons across different action concepts. 521 Italian verbs and 550 English verbs (i.e., the high-frequency verbal lexicon most likely to be used when referring to action) have been processed (Moneglia et al. 2012).

The corpus-based strategy relied on an induction process that separated the metaphorical and phraseological usages from those strictly referring to physical actions.

IMAGACT only specifies the various possible interpretations of verbs with respect to physical actions, while ignoring the other interpretations. Therefore the possible interpretations of verbs beyond physical actions are not considered and are not represented in the ontology.

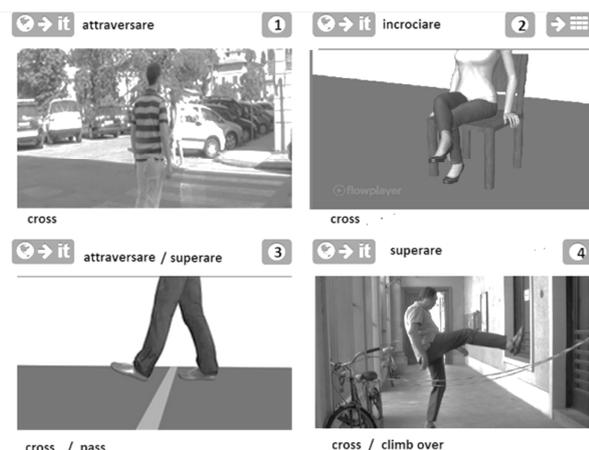


Figure 1. The four action types of the verb *to cross*

The unique visual format of the ontology makes the representation of abstract concepts difficult or impossible. This limitation, however, also constitutes an important added value, which can benefit our knowledge of action verbs in their abstract interpretations and the identification of these meanings within ontologies, as we will show in later sections of this paper.

The capacity to refer to many different physical activities with a single verb belongs to the core of the semantic competence of a language, which has been achieved by mother-tongue speakers during the early phases of their first language acquisition. A speaker cannot assert knowledge of the meaning of *cross* if he is not able to judge that the above events can be the object of its application. At the same time, despite the difference between the different actions represented in each concept, he will also be able to judge that none of them represents the meaning of the verb better than the others and that the

verb is applied in its own meaning in all cases (primary meanings). This is not the case for metaphors, phraseology and abstract meanings.

For instance, the semantic competence of the speaker is not affected if she does not understand the meaning of “John crossed wires with Mary” (idiom) or “John needs to cross to another account” (metaphor). Competent speakers are, on the contrary, able to judge that in these cases the verb is not used in its physical meaning (marked meanings). Nonetheless, roughly half of corpus occurrences of action verbs are not used in their primary, physical meanings, and the use of verbal predication extended from physical meanings is one of the more productive means of reference in natural languages.

This paper describes the infrastructure that has been designed to deal with variations that do not fall within the primary, physical variation of an action verb. It will specifically highlight the benefits that IMAGACT’s representation of the primary meanings through videos brings to the understanding and annotation of secondary meanings.

2. Processing Corpus Occurrences in IMAGACT and the Selection of Marked Variation

The construction of IMAGACT requires the examination and interpretation of verb occurrences in an oral context, which is frequently fragmented and may not provide enough semantic evidence for an immediate interpretation. To this end, the annotation infrastructure allows the annotator to read the context of the verbal occurrence in order to grasp the meaning. The annotator represents the referred meaning with a simple sentence in a standard form for easy processing. This sentence is positively formed, in the third person, present tense, active voice, with the essential arguments of the verb filled. Crucially, along with the standardization, the annotator assigns the occurrence to a “variation class”, either PRIMARY or MARKED (Moneglia et al. 2012).

The decision concerning the status of the occurrence makes use of an operational test roughly derived from Wittgenstein (1953). The occurrence is judged PRIMARY if it is possible to say to somebody who does not know the meaning of the verb V that “the referred action and similar events are what we intend with V”; otherwise the occurrence is MARKED. For instance, the occurrences standardized in “John crosses the finish line”; “John crosses the street” and “John crosses his legs” are assigned to PRIMARY variation, since all can be pointed to explain “what cross means”.

Conversely, the instances standardized as “a thought crossed John’s mind” are not what one uses to instantiate the meaning of *to cross* and therefore have been tagged as MARKED. The annotation of primary versus marked variation has been evaluated at 9.5 K-Cohen agreement (Gagliardi 2014).

The positive selection of occurrences in which verbs refer in their own meaning to physical actions preceded the annotation of action concepts. Only occurrences assigned

to the PRIMARY variation class make up the set of Action Types stored in the ontology. To this end, the standard IMAGACT infrastructure allows clustering of occurrences under prototypes representing the various action concepts, keeping granularity to its minimal level (8.2 K agreement [Gagliardi 2014]). The full annotation process can be found in Moneglia et al. 2012.

Concepts are represented using the universal language of images, which allows the reconciliation, in the IMAGACT ontology, of the types derived from the annotation of different language corpora. 1010 distinct action concepts have been identified and visually represented with prototypical scenes, either animated or filmed (Frontini et al. 2012; Moneglia et al. 2012). The cross-linguistic correspondences of those actions with the verbs that can refer to them in English and Italian have been established in a MySQL database.

38,462 occurrences have been processed in the English corpus and 42,723 in the Italian corpus. Respectively 19,229 and 16,210 (50% and 38%) have been considered marked.

3. Marked Variation Categories

We have established three main categories of marked variation--*metaphor*, *metonymy* and *idiom*--and criteria for creating types within these categories for each verb. The criteria rely heavily on the images that compose the IMAGACT ontology of action and on widely accepted processes of meaning extension in linguistics. Although figurative language is known for its amorphous, subjective nature, we have endeavored to create a standard, justifiable process for determining figurative language types for individual verbs, that we will show in the following sections on the basis of the verbs *to turn* and *to close*.

3.1 Metaphor

The process for identifying a metaphoric type for a verb involves several steps and satisfying several related criteria. First we list all the occurrences of a verb that were labeled as “marked” during the initial corpus annotation of the IMAGACT project. We then use a standard lexicographic procedure of gathering similar usages together. For each group of occurrences that is a potential metaphor, we look for an image or “family” of related images from the IMAGACT ontology to which the occurrences are related. For example, the following list is a sample of one group of corpus occurrences for the verb *to turn*:

John turns to the question of religion

The presenter turns to [the subject of] the book

The colleagues turn to the report

The host turns to the other issues

We have linked this group to the S4 animated video from the IMAGACT ontology shown in Figure 2. The action is of a woman facing straight ahead then turning her head to the right.

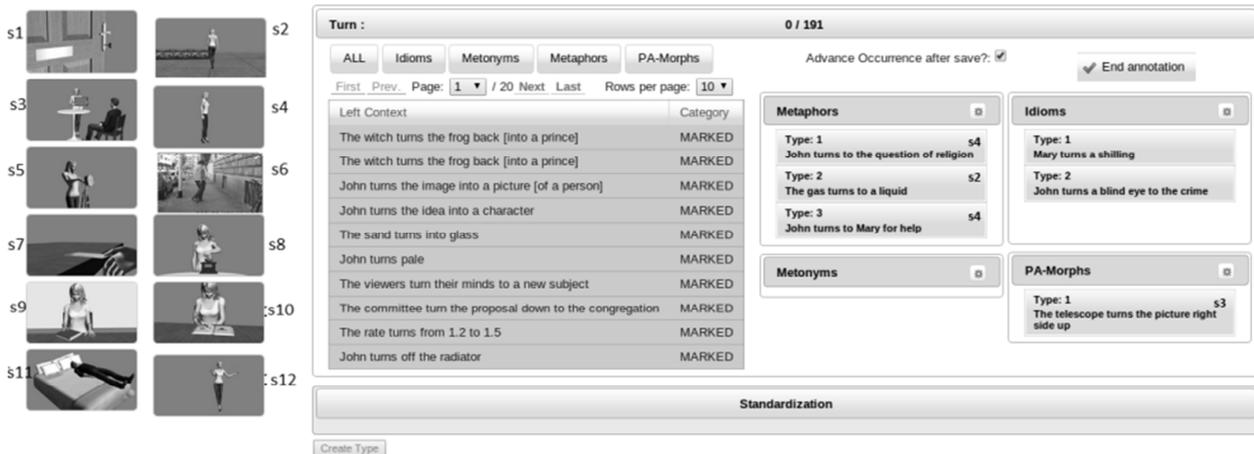


Figure 2. Interface sample for *to turn*

The next step is to identify the property of the action that affords the extension of the verb to a more abstract domain. In this video, the actor turning her head now sees whatever is to her right rather than whatever is directly in front of her. This physical turning of her head can indicate a change in the focus of her attention, say, from a street in front of her to a dog barking on her right. With a metaphorical extension, *to turn* can be used to indicate a change in the focus of one's attention to abstract things as well, such as the question of religion.

One of the most influential theories of metaphor has been that of conceptual metaphor (Lakoff 1987), which posits that a fundamental mechanism of human cognition is the use of a concrete, physical domain to understand a more abstract one. These conceptual metaphors are often revealed in a group of related lexical metaphors. For example, the conceptual metaphor *Life is a Journey* can be seen in the sentences "Mary needs to move on after her divorce" and "the governor ran into a political road block." Where it is possible, we identify the general conceptual metaphor that supports the specific linguistic metaphor in question. Using the list of conceptual metaphors maintained by the University of California, Berkeley, we linked the *turn* metaphor just described to the conceptual metaphors *Change is Motion* and *Ideas are Locations*. Thus, a person facing one location (idea) can turn to face another, indicating a change in her attention from one idea to another.

As with the identification of primary, physical types in IMAGACT, we use equivalent verbs to help distinguish metaphorical types. For the marked variation, we distinguish between equivalent verbs that are used in their primary, or non-figurative, meaning and equivalent verbs that are used in a marked or figurative sense. For example, the verb *shift* has been identified as a verb that can be used in the same situations as *turn* in "John turns to the question of religion." Both of these verbs are used metaphorically in this situation, with the same metaphorical meaning. This match is relevant for an ontology of abstract concepts and corresponds to action concepts in the IMAGACT database.

However, the key means of distinguishing types within the category of metaphor are the links to the action

concepts they derive from and the descriptions of the relevant properties that license the metaphorical extensions. Often, links to different action concepts are enough to distinguish two marked types of a verb. For example, "John turns to the question of religion" is linked to type S4, as described above. Another very common metaphor for the verb *to turn* refers to a change of state, such as "the witch turns the frog back into a prince" or "the gas turns to a liquid". The metaphor is linked to the action concept represented by the video in S2. As part of the conceptual metaphor *Change of State is Change of Direction*, the linguistic metaphor for *turn* in this case uses the property of moving in a new direction from a different action concept and image than the previous *turn* metaphor.

Sometimes two or more metaphors derive from the same action concept but rely on different properties of that concept. Another metaphor of *to turn* links to the S4 image in Figure 2: "John turns to Mary for answers" or "Mary turns to a psychiatrist". In this case, the reorientation of the actor's head indicates an appeal for interaction rather than a change in the focus of his attention. Identifying the prototype related to the metaphor helps in understanding the properties that license the metaphoric extension.

3.2 Metonymy

Metonymy is a less studied phenomenon than metaphor, especially as it pertains to verbs. However, the corpus data we have gathered suggests that it is a necessary category to fully account for the marked variation of certain verbs. For our purposes, we have defined verb metonymy as the use of one action or event to represent a sequence or set of events of which it is a part. For example, many occurrences of *to close* in our English corpus follow the form of "John closed the pub" and "The management closed the factory." This usage of *close* does not follow the process of metaphorical extension, in which an abstract domain is being understood using properties from a physical one. There are actual actions of closing involved in the situations described by these sentences. When John closes the pub, he does indeed close the door. He probably also takes the cash from the register, turns off the lights, and locks the door as he leaves. This is not a physical domain being used to understand an abstract one,

but one action in a sequence of events being used to represent the whole sequence (Goossens 1995).

Complicating the situation, the events in such a sequence are not always all physical actions. “The management closed the plant” probably is also meant to include the decision to end production at the plant, as well as the action of closing and locking the doors. For our purposes, as long as part of the whole event can be described using the verb in its physical sense, we have categorized that type as a metonymic one.¹

For this category, we also link the type to an image from the action ontology. The type of *close* described previously is linked to the video in Figure 3. We also identify one or more equivalent verbs. As with metaphor, where the equivalent verbs are usually other verbs used in a metaphorical way, the equivalent verbs for metonymic types are often other verbs being used metonymically. For example, *shut* is the equivalent verb for this type.

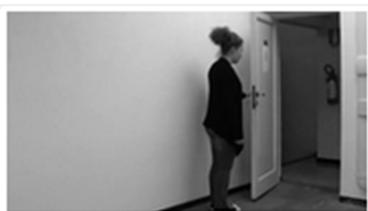


Figure 3. Action type for *to close*

3.3 Idiom

We use a standard definition of idiom: a fixed phrase whose meaning cannot be deduced by combining the meanings of the individual words in the phrase. Because idioms are usually language specific, we have not attempted to link any idioms to the language-independent action concepts in IMAGACT. Instead, we identify an equivalent verb, along with a specific synset in WordNet. For instance, we identify the idiom “turn a deaf ear to” with the equivalent verb *ignore*, connected to the WordNet synset [*neglect, ignore, disregard*].

4. Ongoing Work

We have tested our categories and criteria against the full set of corpus occurrences for five verbs (*turn, cross, pull, close, combine*), creating types to account for all the occurrences. Although this exercise has largely supported the applicability of our schema, it has also raised some questions that we are still in the process of resolving. For some highly frequent verbs, like *to turn*, we find a few, very common marked types. For others, like *to pull*, we find a myriad of different marked types, many of which occur only once or twice in the corpus. How to efficiently account for these rare types remains an open question.

We have also discovered verbs with marked usages that

¹In some cases, a metonymic use of a verb seems to have been further extended into a metaphor. Rather than create a complex annotation scheme where categories can interact, we have provisionally decided to treat these as metaphors.

do not seem fit into any of our three categories, such as *Mary received the wire transfer*. In these cases the verbs appear to have the same meaning as one of the primary, physical types for that verb, but to be acting on objects that are not strictly physical. We are in the process of evaluating a fourth type to account for these usages.

We plan to evaluate further our marked categories and methodology for type creation by annotating the full set of corpus occurrences for a larger set of action verbs from the IMAGACT ontology, a set that includes verbs taken from each of the upper level nodes of the ontology. Based on the results, we will finalize the annotation interface, then use it to process all of the marked occurrences identified by the original IMAGACT annotation. We anticipate supplementary annotation to account for thematic roles and the possible regularities among types that they may reveal (Brown & Palmer 2012). We expect this work to lead to a rich study of the relation between the marked and primary types of high-frequency verbs.

5. References

- Conceptual Metaphor Home Page.
<http://www.lang.osaka-u.ac.jp/~sugimoto/MasterMetaphorList/MetaphorHome.html>
- IMAGACT. <http://www.imagact.it>
- Brown S. W., & Palmer M. (2012). Semantic annotation of metaphorical verbs with VerbNet. Interoperable Semantic Annotation (ISA-8), 8th Joint ISO-ACL/SIGSEM Workshop, Pisa, October.
- Frontini, F., De Felice, I., Khan, F., Russo, I., Monachini, M., Gagliardi, G. & Panunzi, A. (2012). Verb interpretation for basic action types: Annotation, ontology induction and creation of prototypical scenes. CogAlex-III Workshop as part of the COLING 2012 conference. Mumbai (India), December.
- Gagliardi, G. (2014). Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnostica del “mild cognitive impairment” (MCI). PhD Dissertation, University of Florence.
- Goossens, L. (1995). Metaphonymy: The interaction of metaphor and metonymy in expressions of linguistic action. In L.Goossens, P. Pauwels, B. Rudzka-Ostyn, A. Simon-Vanderbergen & J. Vanparys (eds), *By Word of Mouth*. Amsterdam: John Benjamins, pp. 159--174.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Moneglia, M. (in press). Natural language ontology of action. A gap with huge consequences for natural language understanding and machine translation. In Z. Vetulani, J. Mariani (eds), *Post LTC 2011 LNAI*, Berlin: Springer.
- Moneglia, M., Gagliardi, G., Panunzi, A., Frontini, F., Russo, I. & Monachini, M. (2012). IMAGACT: Deriving an action ontology from spoken corpora. 8th Joint ISO-ACL/SIGSEM Workshop, Pisa, October.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.