Finding metaphors at Wikipedia: a frame semantic study

Figueiro Teixeira Lilian

Keio University

lilifiteixeira@gmail.com

1 Introduction

Wikipedia, a collaborative encyclopedia, has been used for different tasks related to language processing, such as semantic relation extraction (Nakayama et al., 2008)[4], music information retrieval (Sordo et al., 2012)[6] and also comprehension and generation of metaphors (Veale and Hao, 2007)[8]. For this preliminary study, we analyzed some metaphoric expressions found at the English version of Wikipedia. The purpose is to verify whether Sullivan's hypothesis about source/target identifications of metaphorical expressions are valid in a specific genre.

2 Sullivan's approach[7]

Sullivan (2013)[7] proposes an innovative approach to the study of the metaphoric language, which combines the concepts of conceptual metaphor, frame semantics and construction grammar. The main features of this approach are described in this section.

In the **Conceptual Metaphor Theory (CMT)** "metaphor is a cognitive process that allows one domain of experience, the target domain, to be reasoned about in terms of another, the source domain" (Lakoff & Johnson, 1980 *apud* Sullivan, 2013, p. 1)[7]. Usually the target domain is an abstract concept, such as TIME, and the source domain is a concrete concept, such as MONEY as, in the famous metaphor TIME IS MONEY. The mappings, the correspondences between the two domains are explained by Sullivan[7] through the analysis of the **semantic frames** evoked by the elements of the metaphoric expressions. Semantic frames are "a script-like conceptual structure that describes a particular type of situation, object, or event and the participants and props involved in it" (Ruppenhofer et al. 2010, p. 5)[5]. The frames for English language that are described on the FrameNet website were used as a reference for Sullivan's study (`https://framenet.icsi.berkeley.edu/fndrupal/`).

Sullivan[7] analyzes different metaphoric **constructions**, pairings of form and meaning, and describes them, through the identification of some patterns in relation to their conceptual autonomy and the mappings between the frames in the different domains. Identifying how the relation of **semantic dependency** of two elements is in a metaphoric expression is crucial to describe the constructional constraints. In any grammatical construction, not just the metaphoric ones, one element is conceptually autonomous and the other is conceptually dependent. In other words, the dependent element requires the conceptualization of other element (the autonomous one) to make sense. (Langacker 1987, 1991, 2002 *apud* Sullivan, p. 8-9)[7]. In Sullivan's study[7], the conceptually dependent element is going to refer to the metaphoric source domain (Croft, 2003 *apud* Sullivan, p. 9)[7], while the autonomous element relates to the target domain.

In order to structure the metaphoric mappings between domains and the target domain, Sullivan (p. 37) points out that the frame relations and frame elements should be preserved. It is easier to understand it by analyzing the following example: *sunny mood*, in which the conceptual metaphor HAPPINESS IS LIGHT (small capital letters are used to indicate concepts) explains the relation between the two domains. The word *sunny* evokes the LIGHT domain, through the LOCATION_OF_LIGHT frame, as it can be seen on Figure 1. There are many frames that are related to the LIGHT domain, but since the metaphoric mappings between the domains are always partial (Lakoff & Johnson, 1980, 1999 *apud* Sullivan, p. 37)[7], the expression highlights only the meaning described by the frame LOCATION_OF_LIGHT. To reinforce that, Sullivan[7] points out that the non-metaphoric use of the lexical item *sunny* is often used before a noun that describes a location, such as in a *sunny street*.
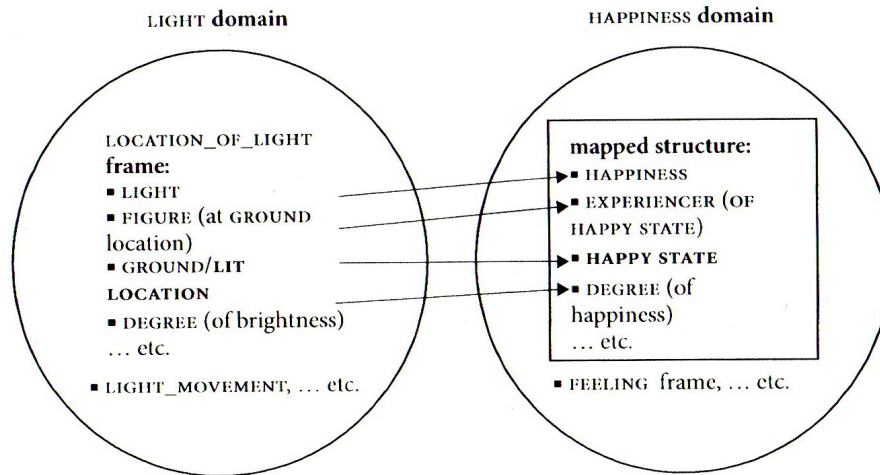
Figure 1: HAPPINESS IS LIGHT Sullivan([7] p. 40)

In order to understand Sullivan's conventions[7], applied in Figure 1, it is necessary to note that the metaphoric domain, mappings, frames, and frame elements are named in SMALL CAPS, while the lexical items are in italics. Besides, the frame evoked is represented as a rectangle inside the domain, which is circular and the arrows represent the correspondences between the relevant frame elements.

For this analysis, I will focus on the following points: (i) the source domain, usually a more concrete concept, is conceptually dependent; while the target domain, usually more abstract, is conceptually autonomous, (ii) the frame structure of the source domain is preserved in the target domain.

3 Example of analysis applying Sullivan's approach [7]

Starting our analysis with the expression *Queen of J-Pop*, we could say that *queen* evokes the LEADERSHIP frame, which is part of the TERRITORY domain, and *J-Pop* evokes the MUSIC domain, being a kind of music. According to Sullivan's approach[7], it could be represented in the following way:
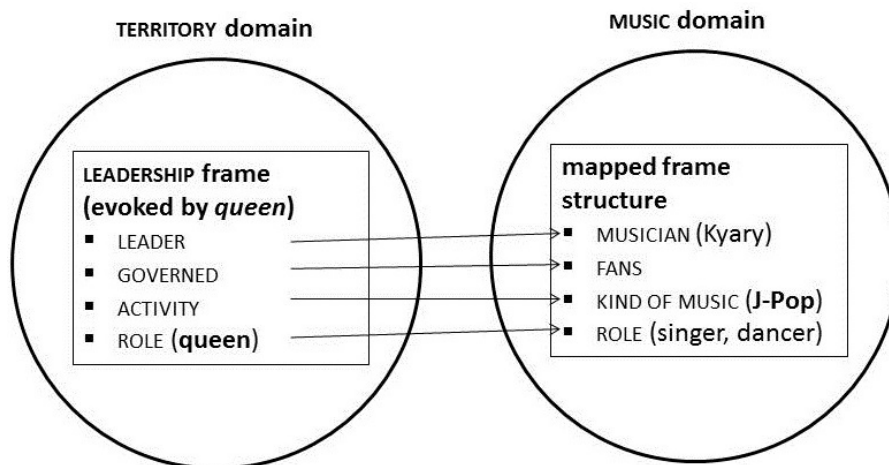

Figure 2: A FIELD (MUSIC) IS A TERRITORY AND THE MUSICIAN IS A LEADER [a]

---

[a]The frames and frame elements presented in the analysis part of this paper are the according to the FrameNet website

In the expression *Queen of J-Pop*, *queen* is a relational noun, because a queen usually rules a territory or is in charge of a specific activity. In *Queen of J-Pop*, according to Sullivan's description of the relational noun in preposition phrases (p. 117), *J-Pop* elaborates the *queen* and *J-Pop* is autonomous in relation to *queen*.

Relational nouns, according to Sullivan (p. 117)[7], evoke a frame in which the second NP fills a specific role, usually elaborating a frame element in the mapped structure. The same occurs here, since

J-Pop is the KIND OF MUSIC that corresponds to the activity in the LEADERSHIP frame. Just as Croft (2003) described the pattern of metaphoric preposition phrase constructions with relational nouns (Sullivan, p. 119)[7], in our example the autonomous element, *J-Pop*, evokes the target domain and the dependent element, *queen*, evokes the source domain.

However, we might question how conventional it really is. For Sullivan ([7], p. 57), a conceptual metaphor should be apparent outside of language and should involve multiple mappings.Other kinds of mappings between the two domains are possible, since we can also say that a musician "commands his fans", while on stage or even outside it. For Kovecses ([2], p. 7), a conceptual metaphor often materializes or is realized not only in language, but also in social-physical practice. As for the musicians, it is known that they have some power over their fans and can even influence them about fashion, politics or other topics. In this sense, the metaphor THE MUSICIAN IS A LEADER expresses how the fans see their idols.

The domains and frames related to second example analyzed, *Internet videos which have gone viral*, are described in Figure 3. This example is an equation, in which we have a copula linking two NPs, *Internet videos* and *viral*. According to Sullivan's approach to equations ([7] p. 105), in this case, *are viruses* is dependent and *Internet video* is autonomous. The subject NP evokes the target domain, *Internet videos*, and the copula-linked XP evokes the source domain.
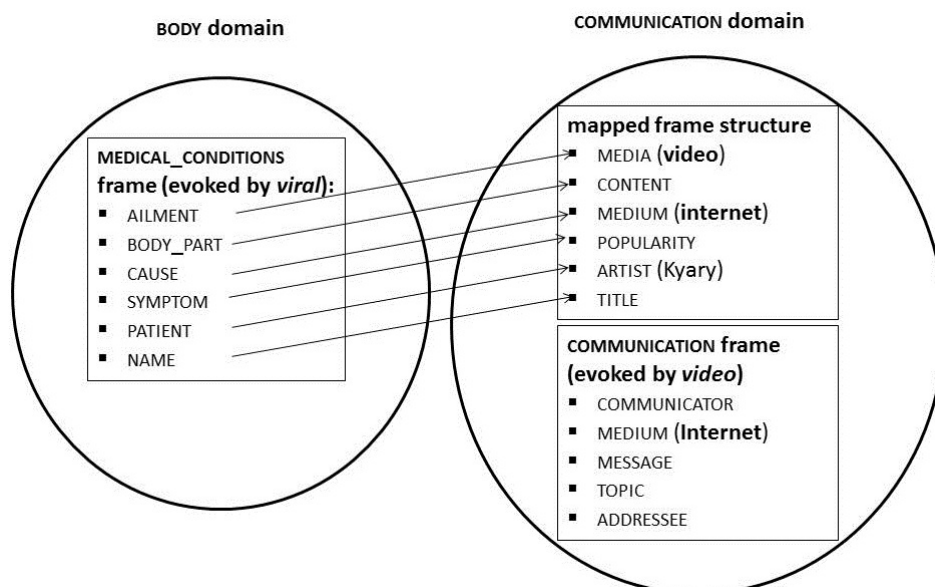

Figure 3: THE CONDUIT METAPHOR (COMMUNICATION IS TRANSFER), IDEAS ARE VIRUS

This kind of construction, equation, as stated by Sullivan ([7] p. 107) often directly evokes one or both domains. Direct evocation happens when an item evokes a domain without profiling a specific frame (Sullivan[7], p. 27), however, as we can see in Figure 3, this expression evokes the frame MEDICAL_CONDITIONS, evoking the domain indirectly.

Sullivan ([7] p. 107) points out that the copula constructions are ideal for expressing image metaphors. Image metaphors differ from most conceptual metaphors in that they do not map the concrete onto the abstract, but rather map one concrete sensory image onto another concrete image. These metaphors are novel and unconventionalized, and it is difficult to identify a target domain. It seems that *"viral videos which have gone viral"* is conventionalized somehow and it is possible to identify THE CONDUIT METAPHOR. However, it is a new form of communication in which the information is shared, by spreading as fast as a virus.

Also Sullivan[7] mentions that image metaphors discourage their expression via predicating modifier and non-copula argument structure constructions. For our example, it would be possible to say "viral video", in which viral could be considered a domain adjective, which is according to Sullivan's description[7] of image metaphor. However, it is also possible to say that "one video is more viral than another", or "the most viral video", in a predicative use.

IDEAS ARE VIRUS can be considered a new subtype of THE CONDUIT METAPHOR, but other uses, such as in the compounds *viral video* or *viral market-*

ing might have lost their metaphoric association, being considered just a term of a specialized language, related to marketing or communication.

4 Concluding remarks and future work

Applying Sullivan's approach to the study of the metaphoric expressions found in Wikipedia facilitated the task of identifying the two domains involved in metaphors, as well as, the elements mapped in the two domains, through the description of the frames. It was also possible to verify how the metaphoric meanings are related to the constructions. In the example about the word *viral*, different expressions, such as *the video has gone viral* and *viral videos* could be interpreted differently, reflecting how a metaphoric expression created to explain a concept in a specific area can acquire a wider meaning when used by non-specialists, as suggested by Deignan et al. [1]. In relation to the example *Queen of J-Pop*, the conventionality of the metaphor could be verified if we consider it a materialization of how the society or a group of people behave [2], reflecting in this case a specific concept related to music.

Considering the findings reported here, we intend to analyze other Wikipedia articles about Japanese music, in order to identify other metaphoric expressions. Since the metaphors can have different meanings according to different genres where they are used, we consider the possibility of using Wikipedia, as a reference corpus to compare it to a more specialized corpus on Japanese music.

Acknowledgements

References

[1] Alice Deignan, Jeannette Littlemore, and Elena Semino. *Figurative Language, Genre and Register.* Cambridge, Cambridge, first edition, 2013.

[2] Zoltán Kövecses. *Metaphor in Culture: universality and variation.* Cambridge, New York, first edition, 2005.

[3] George Lakoff, Jane Espenson, and Alan Schwartz. Master Metaphor List: second draft copy. Technical report, Cognitive Linguistics Group, University of California, Berkeley, 1991.

[4] Kotaro Nakayama, Takahiro Hara, and Shojiro Nishio. Wikipedia Link Structure and Text Mining for Semantic Relation Extraction: towards a huge scale Global Web Ontology. In Stephan Bloehdorn, Marko Grobelnik, Peter Mika, and Thanh Tran Douc, editors, *Proceedings of the Workshop on Semantic Search (SemSearch 2008) at the 5th European Semantic Web Conference (ESWC 2008)*, pages 59–73, Tenerife, Spain, 2008.

[5] Josef. Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. *FrameNet II : Extended Theory and Practice.* International Computer Science Institute, Berkeley, 2010.

[6] Mohamed Sordo, Joan Serrà, Gopala K. Koduri, and Xavier Serra. A Method for Extracting Semantic Information from On-line Art Music Discussion Forums. In *Proceedings of the 2nd Comp-Music Workshop*, pages 55–60, Istanbul, 2012. Universitat Pompeu Fabra.

[7] Karen Sullivan. *Frames and Constructions in Metaphoric Language.* John Benjamins, Amsterdam, first edition, 2013.

[8] Tony Veale and Yanfen Hao. Comprehending and Generating Apt Metaphors: a Web-driven, Case-based Approach to figurative language. In Robert C. Holte and Adele Howe, editors, *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*, pages 1471–1476, Menlo Park, 2007. AAAI Press.