

Building a Polish Corpus of Synesthetic Metaphors

Magdalena Zawisławska (University of Warsaw, Poland)

Introduction

The research project Synamet described in this paper aims at creating a semantically and grammatically annotated corpus of Polish synesthetic metaphors. The paper outlines the procedure followed in the project, the main problems with the metaphor annotation, and the preliminary results.

Synesthetic metaphor is when a perceptual sensation (e.g. olfactory) is described by lexemes that primarily activate another sense (e.g. taste), cf. *sweet aroma*. A broader definition includes the mapping of perception experience on more abstract domains, cf. *sweet love*. Synesthetic metaphors can provide valuable material for preliminary research because their common use in language guarantees that the material collected will be sufficiently rich and varied (which should help later to broaden the analysis to include other types of metaphor). In addition, restricting the research field to the domain of sensory perception will ensure the precise indication of the research scope.

The most important outcomes of the Synamet project are: creation of a valuable source of linguistic data, evaluation of existing models of metaphor and models of synesthesia, and so an important contribution to the redefinition of figurative language.

Method

Most recent corpus-based studies utilize the CMT, formulated by G. Lakoff and M. Johnson (1988). The authors view metaphor as a primarily conceptual phenomenon consisting of mapping across domains (from the source domain onto the target domain). Recently, in some works on metaphor, a 'domain' has been replaced with the concept of frame (Dancygier and Sweetser 2014). Likewise, in the project *MetaNet: A Multilingual Metaphor Repository* two different methodologies are employed: the CMT as well as the FrameNet ontology.

Fillmore posits that the meaning of lexical units, phrases, grammatical and syntactic constructions resides in frames—schematic phenomena such as our beliefs, experiences or typical actions (Fillmore 1982). The metaphorization process can be described as *frame shifting*. Coulson (2001) defines this phenomenon as a "semantic reanalysis process that reorganizes existing information into a new frame." It means that some elements of a frame evoking specific sensations (e.g. smell) may become reorganized under the influence of a lexeme activating a frame of some other sensory perception (e.g. hearing). See, for example:

*Dochodzi **zapach** delikatnych kwiatów, szyprowy **oddech** mchu, a baza **mruczy** rozkosznie delikatnym piżmem.*

'The **smell** of delicate flowers is drawing near, the chypre **breath** of moss, and the base note **purrs** contentedly with delicate musk.'

In the Synamet project, frames have been built up from scratch. The existing FrameNets are not suited to this purpose as there are too many semantic and grammatical differences between Polish and English. In the Synamet corpus, the frames and their elements are adjusted to the texts analyzed—that is, the frame coordinator adds new frames or their elements when the annotators signal that such modification is needed. At present, the corpus tool ATOS contains 6 perceptual frames (SIGHT, HEARING, TOUCH, SMELL, TASTE and MULTIMODAL PERCEPTION—for a sensation that activates several senses, e.g. weight or consistency), and 55 non-perceptual frames (e.g. MAN, ARCHITECTURE, WILD ANIMAL, DOMESTIC ANIMAL, PLANT, SPACE, TIME, ART, SOCIETY, ARMY, HAZARD, etc.). The frames in the project are understood as an analytical tool, not as real conceptual knowledge units, and as they are derived from linguistics data, they are not universal. Every frame element is marked with a typical lexical example, e.g. MAN/emotion (*anger*), WILD ANIMAL/part of animal (*claw*).

Material used in the corpus and the tool for analysis

According to Wering *et al.* (2006), a metaphor is synesthetic only when its source domain pertains to perception (visual, auditory, olfactory, tactile, or gustatory). If the target domain does not evoke perception, we can talk of a *weak synesthetic metaphor*. If both the source and the target domain evoke perception, we are dealing with a *strong synesthetic metaphor* (Werning, Fleischhauer, Beşeoğlu 2006). In Synamet both types of metaphors are annotated.

Since metaphors basically work within a context, the annotators analyze whole texts. The texts are excerpted from blogs devoted to perfume (SMELL), wine, beer, cigars, Yerba Mate, tea or coffee (TASTE, SMELL, SIGHT), as well as culinary blogs (TASTE, SIGHT), music blogs (HEARING), art blogs (SIGHT), massage and wellness blogs (TOUCH).

For the analysis, a dedicated tool called ATOS is used (Annotation Tool of Synesthetic Metaphor). The procedure of annotation includes:

- 1) Extraction of a metaphorical unit, e.g.: *tannin smooth*.
- 2) Correction of the text phrase (if needed), e.g.: *tannin is smooth*.
- 3) Definition of the referent, e.g.: taste (of a wine).
- 4) Description of the phrase type: nominal predicate.
- 5) Selection of the metaphor type: strong (because both frames are perceptual, (see Werning *et al.* 2006).
- 6) Selection of metaphor category: simple synesthesia.
- 7) Definition of the semantic head of the phrase: *tannin*.
- 8) Description of the source frame: TOUCH.
- 9) Selection of the source frame element: TEXTURE.
- 10) Description of the source frame evoking expression: *smooth*, part of speech: adjective.
- 11) Description of the target frame: TASTE.
- 12) Selection of the target frame element: TASTE COMPONENT.
- 13) Description of the target frame evoking expression: *tannin*, part of speech: noun.

Preliminary results of annotation

The statistics produced for 2915 metaphorical units (42% of the planned corpus) show some interesting features of synesthetic metaphors in Polish. The most frequent source perceptual frames are SIGHT (445 units) and TOUCH (297 units). The result is quite concerning, because in all models of synesthesia (Ullman 1962, Judycka 1963, Williams 1976, Classens 1993, Rogowska 2007) the most basic sense is touch.

Table 1. Source perceptual frames

FRAME	UNITS
WZROK (SIGHT)	445
DOTYK (TOUCH)	297
SŁUCH (HEARING)	242
PERCEPCJA ZŁOŻONA (MULTIMODAL PERCEPTION)	227
SMAK (TASTE)	181
ZAPACH (SMELL)	12

Less surprising are the most frequent target perceptual frames: in this case, the ultimate recipient is SMELL (1519). From the diachronic perspective, in Polish all olfactory expressions are metaphorical (see Judycka 1963).

Table 2. Target perceptual frames

FRAME	UNITS
ZAPACH (SMELL)	1519
SŁUCH (HEARING)	524
SMAK (TASTE)	489
WZROK (SIGHT)	72
DOTYK (TOUCH)	10
PERCEPCJA ZŁOŻONA (MULTIMODAL PERCEPTION)	5

The most typical pairs of perceptual frames are: HEARING -> SMELL (203), SIGHT -> SMELL (184), TOUCH -> SMELL (155) and TASTE -> SMELL (151). These results are also contrary to the existing models of synesthesia. Judycka (1963) claims that synesthetic metaphors involving hearing and smell are very rare in Polish. She also contends that primarily visual expressions are used mostly for description of various sound features (e.g. *purity, spaciousness of a tone*).

Table 3. Pairs of source and target perceptual frames

SOURCE FRAME	TARGET FRAME	UNITS
SŁUCH (HEARING)	ZAPACH (SMELL)	203
WZROK (SIGHT)	ZAPACH (SMELL)	184
DOTYK (TOUCH)	ZAPACH (SMELL)	155
SMAK (TASTE)	ZAPACH (SMELL)	151
WZROK (SIGHT)	SŁUCH (HEARING)	97
PERCEPCJA ZŁOŻONA (MULTIMODAL PERCEPTION)	ZAPACH (SMELL)	85
WZROK (SIGHT)	SMAK (TASTE)	75

PERCEPCJA ZŁOŻONA (MULTIMODAL PERCEPTION)	SMAK (TASTE)	58
DOTYK (TOUCH)	SMAK (TASTE)	53
DOTYK (TOUCH)	SŁUCH (TASTE)	39
PERCEPCJA ZŁOŻONA (MULTIMODAL PERCEPTION)	SŁUCH (HEARING)	36

In the case of non-perceptual source frames, the most common are personification and ontological metaphors. There are also quite frequent metaphors with the following target frames: CLOTHES, ARCHITECTURE, SPACE and PLANT. The unexpected, but fairly frequent source frames, are those which are very abstract, such as PHYSICAL QUANTITIES, ABSTRACT IDEAS or LANGUAGE.

Table 4. Source non-perceptual frames

FRAME	UNITS
CZŁOWIEK (MAN)	557
RZECZ (THING)	137
UBRANIE (CLOTHES)	68
ARCHITEKTURA (ARCHITECTURE)	62
PRZESTRZEŃ (SPACE)	56
ROŚLINA (PLANT)	52
SZTUKA (ART)	38
SPOŁECZEŃSTWO (SOCIETY)	37
ZWIERZĘ DZIKIE (WILD ANIMAL)	31
WIELKOŚCI FIZYCZNE (PHYSICAL QUANTITIES)	30
POGODA (WEATHER)	29
ŻYWIOŁY (ELEMENTS)	29
ZDARZENIE/AKCJA (EVENT/ACTION)	21
POJĘCIA ABSTRAKCYJNE (ABSTRACT IDEAS)	20
WOJSKO (ARMY)	19
DOM (HOME)	18
PODRÓŻ (TRAVEL)	16
AKWEN (BASIN)	15
JĘZYK (LANGUAGE)	15
ZWIERZĘ DOMOWE (DOMESTIC ANIMAL)	15

The pairs of non-perceptual source frame and perceptual target frames show that the most personified senses are SMELL (283), TASTE (169) and HEARING (88). In turn, ontological metaphors are typical for HEARING (88) and SMELL (58).

Table 5. Pairs of non-perceptual and perceptual frames

SOURCE FRAME	TARGET FRAME	UNITS
CZŁOWIEK (MAN)	ZAPACH (SMELL)	283
CZŁOWIEK (MAN)	SMAK (TASTE)	169
CZŁOWIEK (MAN)	SŁUCH (HEARING)	88
RZECZ (THING)	SŁUCH (HEARING)	58

SOURCE FRAME	TARGET FRAME	UNITS
RZECZ (THING)	ZAPACH (SMELL)	58
ARCHITEKTURA (ARCHITECTURE)	ZAPACH (SMELL)	45
ROŚLINA (PLANT)	ZAPACH (SMELL)	38
PRZESTRZEŃ (SPACE)	SŁUCH (HEARING)	28
UBRANIE (CLOTHES)	ZAPACH (SMELL)	28
PRZESTRZEŃ (SPACE)	ZAPACH (SMELL)	23
SZTUKA (ART)	ZAPACH (SMELL)	21
ZDARZENIE/AKCJA (EVENT/ACTION)	ZAPACH (SMELL)	21
POGODA (WEATHER)	ZAPACH (SMELL)	20
ZWIERZĘ DZIKIE (WILD ANIMAL)	ZAPACH (SMELL)	20

The most frequent source frame elements used in the synesthetic metaphors analyzed are: type of taste (TASTE FRAME), e.g. *sweet, bitter* (211), element of tone (HEARING FRAME), e.g. *note* (205), weight (MULTIMODAL PERCEPTION FRAME), e.g. *light, heavy* (196), temperature (TOUCH FRAME), e.g. *cold, warm* (173), part of the body (MAN FRAME), e.g. *heart, head* (127).

Acknowledgments

This paper is funded by the National Science Centre in Poland under the project no. 2014/15/B/HS2/00182 titled: *Synamet – the Microcorpus of Synaesthetic Metaphors. Towards a Formal Description and Efficient Methods of Analysis of Metaphors in Discourse*.

References

- Classens, C. (1993). *Worlds of Sense: Exploring the Senses in History and Across Cultures*. London and New York : Routledge.
- Coulson, S. (2001). *Semantic Leaps. Frame-Shifting and Conceptual Blending in Meaning*. Cambridge University Press, Cambridge.
- Dancygier, B. and Sweetser, E. (2014). *Figurative Language*. Cambridge University Press, Cambridge.
- Fillmore, C. (1982). Frame Semantics. *Linguistics in the Morning Calm*. The Linguistic Society of Korea (eds.). Hanshin Publishing Co., Seoul, 111-137.
- Judycka, I. (1963). Synestezja w rozwoju znaczeniowym wyrazów. *Prace Filologiczne*, XVIII, 59-78.
- Lakoff, G., Johnson, M. (1988). *Metafory w naszym życiu*. Warszawa: PIW.
- Rogowska, A. (2007). *Synestezja*. Opole: Oficyna Wydawnicza.
- Ullman, S. (1962). *Semantics: an introduction to the science of meaning*. Oxford: Blackwell.
- Werning, M., Fleischhauer, J., Beseoglu, H. (2006). The cognitive accessibility of synaesthetic metaphors. (Eds.) R. Sun & N. Miyake. *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society* (pp. 2365–70). London: Lawrence Erlbaum Associates.
- Williams, J. M. (1976). Synaesthetic adjectives: A possible law of semantic change. *Language*, 52(2), 461–478.