

Magdalena Zawistawska
University of Warsaw
Warsaw, Poland

Marta Falkowska
University of Warsaw
Warsaw, Poland

Maciej Ogrodniczuk
Institute of Computer Science, Polish Academy of Sciences
Warsaw, Poland

VERBAL SYNAESTHESIA IN THE POLISH CORPUS OF SYNAESTHETIC METAPHORS

ABSTRACT

This paper offers the use of figurative language found in online blogs and used as source material for the research project *Synamet*, the aim of which is to create a semantically and grammatically annotated corpus of synaesthetic metaphors for Polish. A particular emphasis of the proposed article is laid upon the verbal synaesthesia found in the corpus. It is argued that the analysis of synaesthetic metaphors found in the material depicts some limitations of the Conceptual Metaphor Theory as proposed by Lakoff & Johnson (2008 /1980/) when applied to discourse analysis. Therefore, more efficient methods of describing figurative expressions should be elaborated.

1. INTRODUCTION

This paper outlines the use of figurative language of online blogs used as source material for the research project *Synamet*. This project aims to create a semantically and grammatically annotated corpus of synaesthetic metaphors for Polish. In the paper we present the main findings on verbal synaesthesia in the corpus — that is, a new model of verbal synaesthesia, the frequency of source and target domains (perceptual and non-perceptual) in synaesthetic metaphors, and the lexemes most often used in these blogs. We also argue that the analysis of synaesthetic metaphors found in our material depicts some limitations of the Conceptual Metaphor Theory (CMT, Lakoff & Johnson 2008 /1980/) when applied to discourse analysis, and that it is necessary to elaborate more efficient methods of describing figurative expressions.

2. SYNAESTHETIC METAPHORS

Our project deals solely with linguistic synaesthetic metaphors (verbal synaesthesia), and not with synaesthesia as a psychological phenomenon. Judycka (1963) reaches the same conclusion as Doroszewski (1927: 30), who claims that synaesthesia in language is a semantic change caused by intertwining various sensory stimuli. Cacciari (2008) also claims that synaesthetic metaphors in language are in general a result of a real psychological phenomenon called synaesthesia. Nevertheless, the latest research weighs against those statements. Synaesthesia as a psychological phenomenon unites simple perceptual sensations on a very basic level of information processing, whereas synaesthetic metaphors connect much more complex mental images. While synaesthesia is an involuntary and automatic experience, synaesthetic metaphors in discourse are as a rule created in a conscious and intentional way. Also, synaesthetic percepts do not change over time and are independent of various outside factors, whereas synaesthetic metaphors significantly evolve over time and are dependent on culture, emotions, context, and knowledge of the world. Synaesthetes are unable to explain why they have synaesthetic associations, whereas the author of a synaesthetic metaphor is aware of the metaphor's origin (see Rogowska 2007: 25). Buttler (1978: 115) defines synaesthetic metaphor as “using one name for various types of sensory perception,” e.g. olfactory perception is described by lexemes that primarily activate another sense (e.g. taste), cf. *sweet aroma*. The synaesthetic metaphor category embraces various subtypes. Judycka (1963: 59–60) cites the taxonomy proposed by Jaensch (1929), who distinguished synaesthesia in the narrow sense (expressions denoting only the perceptual sphere) and synaesthesia in the broad sense (emotional synaesthesia, such as e.g. *dark despair*, conceptual symbols, and the so-called complex synaesthesia). Judycka herself discriminates between *word*

Key words

synaesthetic metaphors, *Synamet*, discourse analysis, online blogs, Conceptual Metaphor Theory

synaesthesia (i.e. etymological, entrenched), e.g. lęk ‘fear’ (Proto-Slavic *lęk-ŋ ‘I bend’), and *phraseological synaesthesia* (e.g. *sweet sound*). She also makes a distinction between simple and complex synaesthesia, the latter category consists in merging the sensations coming from different perceptual domains, e.g. thin/lean, fat/greasy. According to Werning et al. (2006), a metaphor is synaesthetic only when one domain pertains to perception (visual, auditory, olfactory, tactile, or gustatory). If only one of the domains evokes perception, we can talk of a *weak synaesthetic metaphor*. If both the source and the target domain evoke perception, it is a *strong synaesthetic metaphor*. In order to collect the most varied material possible, the project employed the broadest sense of the term *synaesthetic metaphor*. In *Synamet*, both types of metaphors (strong and weak) were annotated. However, the word *synaesthesia* (as Judycka refers to it) was excluded, since expressions of this type require etymological analysis.

Synaesthetic metaphor has been studied not only by linguists, but also by psychologists and anthropologists (Ullmann 1962; Williams 1976; Classens 1993). It appears to be a quite universal phenomenon for natural languages. Synaesthetic metaphors are a valuable field for research because their ubiquity in language guarantees that the collected material will be rich and varied enough (which should later help to broaden the analysis to include other types of metaphor).

3. SOURCES OF THE SYNAMET CORPUS

We believe that metaphor is primarily a pragmatic phenomenon that depends heavily on broad context. We agree with Jensen (2017: 262) that “it is vital (...) to attempt to capture metaphoricity in its natural surroundings.” Therefore, it was essential for us to focus on analyzing authentic texts instead of fabricated ones or highly lexicalized idioms. In our opinion, the best source of such material are blogs. The *Synamet* corpus contains texts from blogs devoted to fields where, we presumed, synaesthetic metaphors were most likely to be found, e.g. blogs devoted to perfume (SMELL), wine, beer, Yerba Mate, or coffee (TASTE, SMELL, VISION), as well as culinary blogs (TASTE, VISION), music blogs (HEARING), art blogs (VISION), massage and wellness blogs (TOUCH). We analyzed entire texts from the chosen blogs (that is blog entries), and not just excerpts.

4. METHOD

4.1. BASIC THEORETICAL ASSUMPTIONS

The analytical approach adopted for *Synamet* draws on both frame semantics (Fillmore 1982) and Cognitive Metaphor Theory (CMT), formulated by

Lakoff&Johnson (2008 /1980/). Within CMT, metaphor is viewed as a primarily conceptual phenomenon consisting in mapping across cognitive domains (from the source domain onto the target domain), e.g. LOVE IS FIRE. Although we have chosen CMT for our project purposes, we are aware of the disadvantages of this method in discourse analysis. First of all, it has not been precisely stated what the term *conceptual domain* stands for, or how the domain's structure is supposed to be reconstructed. Moreover, the metaphors in CMT are presented as very general conceptual schemata X is Y, which results in overlooking the metaphorical expression's semantic and grammatical properties. Neither does the CMT sufficiently account for the cultural factors in metaphor production (cf. Gibbs 2017: 9), which is essential in the case of verbal synaesthesia. According to Hellesten:

[T]he distinction between conceptual, implicit metaphors and linguistic expressions has proved to be problematic because it is often impossible to draw lines between the different levels of the same metaphor and thus to organize the specific metaphorical expressions into a conceptual metaphor. (Hellesten 2000: 20)

Therefore, we concentrated primarily on metaphorical expressions in the analyzed texts. We also decided to try to overcome the shortcomings of the CMT by combining it with frame semantics. Instead of the term *domain* we use *frame*, because while the internal structure of domains is not fully clear, frames are described as ordered structures which contain categories (*slots*) and their values (*fillers*). Recently in some works on metaphor, the concept of domain has also been replaced with the concept of frame (Sullivan 2006, 2013; Dancygier & Sweetser 2014). Likewise, in the project *MetaNet: A Multilingual Metaphor Repository*¹, two different methodologies (CMT as well as the FrameNet ontology) are employed.

In modern linguistics, frame semantics (Fillmore 1982) is one of the most prominent theories. Fillmore posits that the meaning of lexical units, phrases, grammatical and syntactic constructions resides in schematic phenomena, such as our beliefs, experiences or typical actions. The metaphorization process can thus be described as frame shifting. According to Sullivan (2006, 2013) and Petersen et al. (2007), we assume that mapping between frames in synaesthetic metaphors is not accidental but depends on the structures of both frames (the source and the target) — that is, the elements and values of the source frame are mapped onto the analogous attributes of the target frame.

In *Synamet*, frames were built up from scratch. The existing Berkeley FrameNet² was not sufficient for our purposes, as there were too

1 <https://metanet.icsi.berkeley.edu/metanet/>

2 <https://framenet.icsi.berkeley.edu/fndrupal/>

many semantic and grammatical differences between Polish and English. In the *Synamet* corpus, the frames and their elements were adjusted to the analyzed texts from the blogs — that is, the frame coordinator added new frames or their elements whenever the annotators signaled that such modifications were needed. At present, there are 6 perceptual frames (VISION, HEARING, TOUCH, SMELL, TASTE and MULTIMODAL PERCEPTION — for sensations that activate several senses, e.g. weight or consistency), and 55 non-perceptual frames (e.g. PERSON, ARCHITECTURE, PLANT, SPACE, TIME, ART, SOCIETY, ARMY, MAGIC, HAZARD, etc.). The frames in this project are understood as analytical tools rather than as real conceptual knowledge units; further, as they were derived from linguistics data, they are not universal. Every frame element is marked with a typical lexical example, e.g. the PERSON frame/EMOTION element (*anger*), the ANIMAL frame/PART OF ANIMAL element (*claw*).

4.2. RECOGNIZING AND ANNOTATING METAPHORICAL EXPRESSIONS IN SYNAMET

There are several different procedures of recognizing metaphors in discourse. For instance, we have at our disposal the *met** system (Fass 1991), procedure called MPA (*Metaphorical Patterns Analysis*) proposed by Stefanowitsch (2006: 65-70), or Krishnakumaran&Zhu's system (2007), which is based on the WordNet ontology, the MIP system (Pragglejaz 2007; Semino 2008), the MIPVU system (Steen et al. 2010), as well as metaphor identification using nouns and verbs clustering (Shutova, Lin & Korhonen 2010). The most widely known metaphor recognition procedure is MIPVU (Steen et al. 2010), which constitutes a modified and elaborated version of the MIP as proposed by the Pragglejaz group (Pragglejaz 2007; Semino 2008: 11-12). Although the MIPVU had been successfully applied to English and other languages, it was modified for the analysis of texts in *Synamet* corpus. This section provides an overview of process of identifying and annotating synaesthetic metaphors in our corpus. A detailed instruction for annotators was prepared, and the analysis procedure was facilitated by a dedicated computer application called ATOS³ (*Annotation Tool of Synaesthetic Metaphor*). First of all, the annotator first had to define referents (which we called *topics*) of the analyzed discourse. The initial step was the most crucial moment in the analysis of a text, since in Polish, which is a highly inflected language, lexemes denoting referents could easily be omitted in a sentence. For example, in the fragment *Encens Mythique d'Orient dość*

3 ATOS is a tool designed specifically for the project by Magdalena Zawisławska (the principal investigator of the *Synamet* project) and Katarzyna Głowińska (the head of the company Lingventa <http://www.lingventa.pl/index.html>).

szybko zmierza do ambrowego akordu bazy (*Encens Mysthique d'Orient makes its way towards the amber chord of the base quite quickly*)⁴, one of the referents (topics) is the perfume *Encens Mysthique d'Orient*. The next step was extraction of the *metaphorical units* (MUs) that denoted the referents. A metaphorical unit is understood as a single phrase, sentence, or text fragment where two different frames are activated and at least one of them is perceptual. In the above example we have three MUs: 1) *Encens Mysthique d'Orient makes its way towards (...) quite quickly* (the SMELL frame and the PERSON frame), 2) *the amber chord* (the SMELL frame and the HEARING frame), 3) *the base* (the SMELL frame and the ARCHITECTURE frame). Afterward, the annotator opened the editor of metaphorical units. The annotator was obliged to describe the phrase type (e.g. *the amber chord* is a nominal phrase), and to select the metaphor type (e.g. *the amber chord* is a strong synaesthetic metaphor because both frames employed in the expression are perceptual). The next step was the description of a source frame (in the example cited above it is the HEARING frame) and its element (in this case—GROUP OF NOTES: *chord*). A target frame and its element were also described (in this example it is the SMELL frame and the element TYPE OF SMELL: *amber*). The annotator also added a grammatical description of the words activating frames (*activators*)—in this case: noun and adjective. Annotators could also indicate borderline metaphors (there was a STORE catalogue for any questionable phenomena in ATOS).

In *Synamet*, we employed a series (cascade) annotation procedure instead of a parallel one since the experiment set up for the Polish Coreference Corpus proved that the outcome of the series annotation is better than the outcome of the parallel annotation (Ogrodniczuk et al. 2015: 89). Therefore, texts in *Synamet* were first analyzed by four annotators⁵ and subsequently by three superannotators who amended the initial annotation and checked the corpus coherence. Although the series annotation was chosen as the main method in *Synamet*, an experiment was performed with a parallel annotation and annotator inter-agreement was tested for a small subset of texts (40 blog entries analyzed independently by two annotators) with the Cohen's κ 'kappa'. Results were as follows: moderate level of agreement for typical metaphors ($\kappa = 0.77$), almost perfect level of agreement for mixed metaphors ($\kappa = 0.95$), a strong level of agreement for entangled metaphors ($\kappa = 0.80$), and almost perfect level of agreement for narrative metaphors ($\kappa = 0.93$). The experiment showed that despite the complex procedure of annotation and problematic data the annotator inter-agreement was high enough to expect that the corpus would be relatively trustworthy.

4 All examples used in the paper are excerpts from the *Synamet*.

5 All annotators and superannotators were Masters or PhDs in linguistics with specialization in semantics.

5. RESULTS

The *Synamet* corpus contains:

1. 684 528 tokens,
2. 1414 annotated texts from blogs (i.e., blog entries),
3. 2597 metaphorical topics (i.e., referents of metaphorical expressions),
4. 15855 activators (words that activate source or target frames), and
5. 9217 grammatically and semantically annotated metaphorical expressions (metaphorical units—MUs).

The metaphorical productivity of blog categories differs, as presented in Table 1. However, the Kendall correlation test shows no statistically important correlation between the number of tokens and the number of metaphor units in each category (p -value = 0.1092, t = 0.309090):

Table 1. Tokens and metaphorical units in categories of texts in *Synamet*

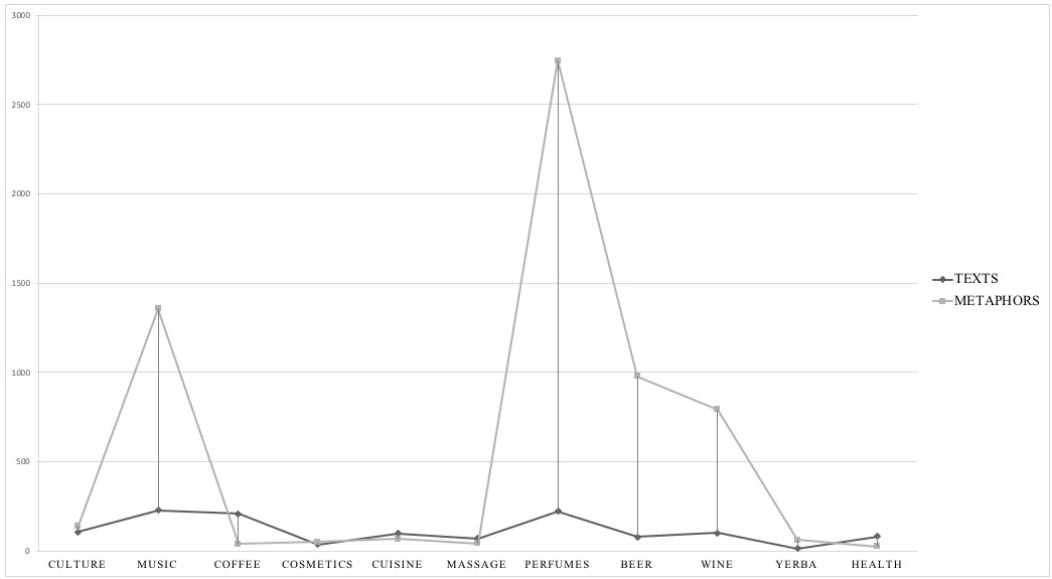
Category	Tokens	MUs
Beer	40 149	1 021
Coffee	90 767	68
Cosmetics	15 437	55
Cuisine	86 966	88
Culture	50 947	145
Massage	80 378	41
Music	127 850	2 291
Perfume	91 328	3 575
Wellness	40 585	49
Wine	54 529	935
Yerba	6 712	53

Figure 1. presents the main categories of the blogs in the corpus and their metaphorical productivity.⁶ The most productive are blogs devoted to PERFUMES, MUSIC, BEER, and WINE. The remaining categories have a relatively smaller contribution to synaesthetic metaphors.

5.1. SOURCE AND TARGET FRAMES IN SYNAESTHETIC METAPHORS

The most frequently used source perceptual frames in the analyzed blogs include the VISION frame, the MULTIMODAL PERCEPTION frame, and the TOUCH frame. The ultimate target frame is the SMELL frame. Table 2 shows source and target frames in the standardized Pearson's residuals ($\chi^2=4594.4$,

⁶ Certainly, the corpus is not fully balanced, but at this stage this is not an objective that we have set for ourselves, since microcorpora are essentially thematically organized.



$df = 5$, $p\text{-value} < 0,001$). All differences between frames are statistically significant, which is presented in Table 2.

Figure 1. Number of texts within thematic categories vs. the categories' metaphorical productivity

Table 3 presents Pearson's standardized residuals of frame pairs in strong synaesthetic metaphors ($\chi^2 = 870.35$, $df = 25$, $p\text{-value} < 0.001$). The most frequent pairs are: HEARING (source) → SMELL (target), VISION (source) → HEARING (target), TASTE (source) → SMELL (target), MULTIMODAL PERCEPTION (source) → TASTE (target).⁷

Table 2. Pearson's residuals of the source and target frames in strong synaesthetic metaphors ($p\text{-values} < 0,001$)

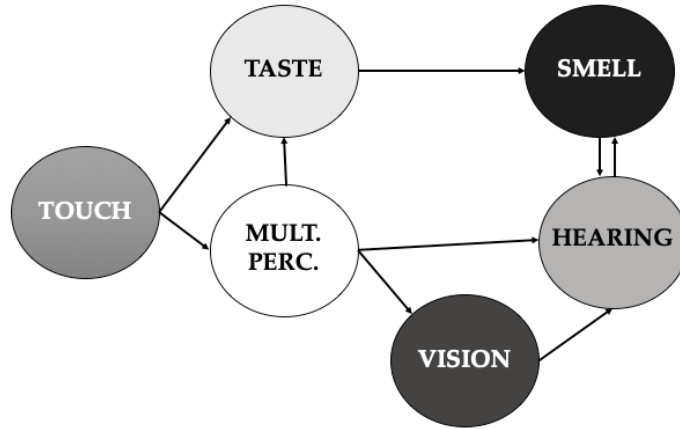
	HEARING	MULTIMODAL PERCEPTION	SMELL	TASTE	TOUCH	VISION
SOURCE FRAME	3.21714	28.31106	-54.04338	-14.28765	14.28765	32.272
TARGET FRAME	-3.21714	-28.31106	54.04338	14.28765	-27.77515	-32.272
Source Target	HEARING	MULT. PERC.	SMELL	TASTE	TOUCH	VISION
HEARING	3.21714	2.644992	4.6927382	-2.25641879	1.699175	13.7583162
MULT. PERC.	-1.471121	28.31106	-0.1209030	-0.87556277	2.586414	0.8689325
SMELL	19.427053	-12.190764	54.04338	11.30972314	-4.557809	-10.8430464
TASTE	-4.802936	9.627941	-1.4921657	14.28765	2.593427	1.1327295
TOUCH	-1.759014	2.457377	-0.1445632	0.01546881	14.28765	0.8377242
VISION	-3.313865	7.055469	2.2409309	-1.60628493	3.542616	32.272

Table 3. Source-and-target pairs of frames in strong synaesthetic metaphors⁷

Figure 2 presents the model of verbal synaesthesia in Synamet:

⁷ Numbers in boldface are statistically significant (for a $p\text{-value} < 0.01$ numbers bigger than 2,58 or smaller than -2,58 are statistically significant).

Figure 2. Model of verbal synaesthesia in Synamet



In weak synaesthetic metaphors, the most frequent pairs of source and target domains include SPORT (source) → TASTE (target), SPACE (source) → HEARING (target), THING (source) → HEARING (target), WEATHER (source) → HEARING (target), ARCHITECTURE (source) → SMELL (target), PLANT (source) → TASTE (target), PERSON (source) → SMELL (target). See Table 4 presenting Pearson’s standardized residuals ($c^2 = 526.32$, $df = 16$, $p\text{-value} < 0.001$):

Source \ Target	HEARING	TASTE	SMELL
ARCHITECTURE	-2.818659	-4.524547	6.266083
ART	-0.4312885	-2.1051118	2.0886845
CLOTHES	1.315318	-1.351310	-0.154747
PERSON	-5.1182337	0.1791543	4.6642333
PLANT	-4.640675	6.251759	-0.640944
SPACE	8.063878	-4.223139	-4.196844
SPORT	-6.691135	16.449416	-6.870617
THING	7.229340	-2.326573	-4.929763
WEATHER	6.544754	-4.912604	-2.218539

Table 4. Source-and-target pairings of frames in weak synaesthetic metaphors (Pearson’s standardized residuals)

The SMELL as a target frame interconnects with the largest and most diverse set of source non-perceptual frames — 44 frames (e.g. PERSON, THING, ARCHITECTURE, PLANT, CLOTHES, ART, WILD ANIMAL, SPACE, WEATHER, ELEMENTS, SOCIETY, LANGUAGE, HOME, ARMY, TIME, WEATHER, BASIN, MAGIC, MACHINE, etc.). The next largest target frame, which receives metaphorical transfer from 37 various non-perceptual source frames, is the HEARING frame (e.g. PERSON, THING, SPACE, TRAVEL, VEHICLE, HEALTH SERVICE). A slightly smaller number — 27 source frames — appear in synaesthetic metaphors with the TASTE as a target frame (e.g. PERSON, SPORT, THING, SOCIETY, ARMY, LANGUAGE). By contrast, the frame VISION as a target frame occurs in metaphors with only 16 different non-perceptual source frames (e.g. PERSON,

PLANT, THING). Frames TOUCH and MULTIMODAL PERCEPTION never appear as target frames in weak synaesthetic metaphors.

Strong synaesthetic metaphors in *Synamet* vary depending on a frame element’s frequency. Some frame elements are selected more often than others. The target elements apparently arise from the main subjects of the blogs, e.g. taste, smell, type of smell, type of taste, or song. More interesting are the elements of the source frames. The broadest set of elements is activated in the VISION frame — 52 different elements, from the HEARING frame — 39 elements, from the TOUCH — 31 elements, from the TASTE frame — 21 elements, and from the MULTIMODAL PERCEPTION frame — 18 elements. Table 5 presents most frequent perceptual frame elements in metaphorical units:

FRAME	ELEMENT	NUMBER OF MUs
HEARING	COMPONENT (e.g. note)	494
MULTIMODAL PERCEPTION	WEIGHT (e.g. light)	444
TOUCH	TEMPERATURE (e.g. hot)	181
VISION	SIZE (e.g. big)	176
TASTE	TYPE (e.g. sweet)	129

Table 5. Most frequent perceptual frame elements in metaphorical expressions in *Synamet*

In synaesthetic metaphors, the most typical pairs of source-and-target frame elements are presented in Table 6:

SOURCE FRAME	ELEMENT	TARGET FRAME	ELEMENT	NUMBER OF MUs
HEARING	COMPONENT (e.g. note)	SMELL	COMPONENT OF SMELL	306
PERSON	BODY PART (e.g. hand)	SMELL	COMPLEX SMELL	159
ARCHITECTURE	PART OF CONSTRUCTION (e.g. stairs)	SMELL	COMPLEX SMELL	107
TASTE	TYPE OF TASTE (e.g. sweet)	SMELL	TYPE OF SMELL	97
TOUCH	TEMPERATURE (e.g. hot)	SMELL	SMELL	96
MULTIMODAL PERCEPTION	WEIGHT	SMELL	COMPONENT OF SMELL	86

Table 6. Most frequent pairings of source-and-target elements

5.2. ACTIVATORS

All activators (i.e. words that primarily activate an element of a frame) in *Synamet* are recorded with respect to two forms: a text form and a base form. Therefore, the base forms of all activators used in synaesthetic metaphors form a mini-dictionary of *Synamet*. There are quite surprising differences between sets of activators interconnected with frames in the corpus. The largest set of lexemes activates the SMELL frame (458) and the HEARING frame (426). The smallest set of lexemes is linked to the MULTIMODAL PERCEPTION frame (142), see Table 7. The perceptual lexemes most often used in the texts gathered in *Synamet* are as follows:

1) in the HEARING frame: *nuta* ‘note’ (nf⁸= 304.005), *kompozycja* ‘composition’ (nf= 111.636), *akord* ‘chord’ (nf= 67.486), *utwór* ‘piece’ (nf= 63.702), *brzmienie* ‘sound’ (nf= 58.656);

2) in the MULTIMODAL PERCEPTION frame: *lekki* ‘light’ (nf= 144.433), *ciężki* ‘heavy’ (nf= 51.718), *lekko* ‘lightly’ (nf= 44.15), *głęboki* ‘deep’ (nf= 35.320), *głębka* ‘depth’ (nf= 17.029);

3) in the SMELL frame: *zapach* ‘smell’ (nf= 247.240), *aromat* ‘aroma’ (nf= 61.810), *perfumy* ‘perfume’ (nf= 61.179), *zapachowy* ‘saporific’ (nf= 29.012), *róża* ‘rose’ (nf= 21.120);

4) in the TASTE frame: *wino* ‘wine’ (nf= 116.682), *słodki* ‘sweet’ (nf= 71.270), *słodycz* ‘sweetness’ (nf= 48.465), *smak* ‘taste’ (nf= 48.565), *gorzyczka* ‘bitter taste’ (nf= 37.842);

5) in the TOUCH frame: *delikatny* ‘fragile’ (nf= 66.855), *ciepły* ‘warm’ (nf= 47.303), *ostry* ‘sharp’ (nf= 42.888), *miękki* ‘soft’ (nf= 20.182), *gładki* ‘smooth’ (nf= 19.552); 6) in the VISION frame: *tło* ‘background’ (nf= 54.872), *wysoki* ‘high’ (nf= 44.150), *czysty* ‘clear’ (nf= 35.950), *barwa* ‘tint’ (nf= 25.859), *pojawiać się* ‘to appear’ (nf=25.859).

Table 7. Activators of perceptual frames in Synamet

FRAME	Adj	Adv	N	V	LEXEMES IN TOTAL
HEARING	86	17	289	34	426
MULTIMODAL PERCEPTION	67	19	35	21	142
SMELL	110	3	336	9	458
TASTE	98	13	210	9	330
TOUCH	88	13	35	44	180
VISION	158	23	122	84	387

5.3. PROBLEMATIC METAPHORS

Since we have chosen CMT as the main method of description, the ATOS tool was designed according to this theory. The ATOS’ editor of metaphorical expression allows annotators to describe typical metaphors, e.g. *Womanity*⁹ *gaśnie powoli* (*Womanity fades slowly*) — SMELL IS LIGHT, or *Tyle innych burgundów cierpi na sklerozę i haluksy* (*So many other burgundies suffer from sclerosis and bunions*) — WINE IS A PERSON. However, during the annotation of texts from blogs, we distinguished three types of metaphorical expressions that cannot be described using the method of CMT (or any other theory of metaphor). The annotation tool needed to be adjusted to capture these new phenomena. The problematic metaphors found in our corpus are mixed metaphors (cf. Kövecses 2016: 3), “entangled” metaphors, and extended (narrative) metaphors (Gibbs 2017: 51; Ritchie 2017).

⁸ Normalized frequency.

⁹ A name of perfume by Thierry Mugler.

5.3.1. MIXED METAPHORS

Mixed metaphors are typical phenomena in blogs. It means that one topic (referent) can be combined with several source frames in one phrase, sentence, or text fragment, e.g. (1):

(1) (...) neroli brzmi tak jak zazwyczaj, czyli głównie zielono i kwiatowo, z lekką nutą octanową (to jaśmin), która wydaje mi się nie na miejscu. Składniki oud — toczony grzybem drewno, pewna smolistość — na chwilę ustępują świergotowi neroli, ale potem łagodnie obejmują przewodnictwo, a neroli dziwnie zaokrągla całość.

(...) **neroli** [topic] **sounds** [HEARING] as usual, that is mainly **greenly** [VISION] and flowery with **a light** [MULTIMODAL PERCEPTION] acetate **note** [HEARING] (it is jasmine) which seems to be out of **place** [SPACE]. **Oud** [topic] components—**wood eaten by fungus** [VISION], some of **pitch-blackness** [VISION]—for a moment give in for a **chirping** [HEARING] of neroli, but later gently **take the leadership** [PERSON], while neroli strangely **rounds off** [VISION] the **whole** [THING].

The main topics of the blogs where we found the biggest number of mixed metaphors are music, wine or beer, and perfumes. The problem of mixed metaphor annotation in *Synamet* is that CMT allows us to describe only two domains (one source and one target) at the same time. Therefore, we had to describe the mixed metaphors in *Synamet* in two steps. First, the mixed metaphor had to be tagged in a text, e.g. (2):

(2) Santo Stefano i Santo Stefano Riserva to wina głębokie, długie, eleganckie, o twardej ramie garbników, a zarazem mięsistej, sutej materii.

Santo Stefano i Santo Stefano Riserva [topic] are deep, long, elegant wines, with a hard frame of tannins and at the same time fleshy, ample fabric.

The next step was to divide the whole mixed metaphor into smaller “X is Y” pieces, e.g.: *deep wines* [MULTIMODAL PERCEPTION → TASTE], *long wines* [VISION → TASTE], *elegant wines* [PERSON → TASTE], *a hard frame of tannins* [ART → TASTE], *fleshy, ample fabric* [TOUCH → TASTE].

5.3.2. ENTANGLED METAPHORS

There is yet another type of metaphor that made annotation quite problematic. We have named it an *entangled metaphor*. In this case the syntax of the metaphorical phrase is contradictory with its semantics (that is a

hypallage¹⁰) and semantically it is a highly compensated layered metaphor very often conflated with metonymy. For example, a sentence *Ta płyta to mroczna podróż przez 20 utworów* (“This record is a dark journey thorough 20 pieces”) might at first sight resemble a simple synaesthetic metaphor. But as we look closer, we can see that the adjective *dark* does not refer to the noun it modifies (*journey*) but rather to the music on the record. The phrase contains a compressed and multilevel metonymic and metaphorical meaning: the record (that is—metonymically—music) is dark [VISION → HEARING] and the record (that is listening to music) is a journey [JOURNEY → HEARING]. In *Synamet*, we chose to annotate such metaphors in three steps. First, the fragment *record is a (...) journey* was described: the word *record* activated the HEARING frame and the element MUSIC, and the word *journey* activated the JOURNEY frame and the element JOURNEY. In the metaphor editor the box “entangled” metaphor was checked and the noun *journey* was marked as the “mediator”, that is an activator which is reorganized by frame shifting—by moving it from the original frame to a new one. The next step was annotation of the phrase *dark journey*. Again, the box “entangled” metaphor in the editor was checked and the noun *journey* was marked as the “mediator”, but this time it was described as an activator of the HEARING frame. And the last step was annotation of the phrase *record is dark (...)*.

5.3.3. EXTENDED METAPHORS

An extended metaphor (or as we call it — *a narrative metaphor*) is “one conceptual metaphor that motivates several related linguistic expressions in the same stretch of discourse” (Gibbs 2017: 47). Ritchie (2017: 242) calls this phenomenon a metaphorical story and defines it as “subcategory of communication in which a vehicle story is at least narrated, such that it expresses something about, and can be mapped onto, a topic story from a totally different domain”, e.g. (3):

(3) **Hedonist Rose** to pełna dziewczęcego uroku towarzyszka, która spędzi z Tobą przy butelce wina cały dzień, od niechcenia rzucając zalotne spojrzenia i chichocząc. Od początku myślisz, że wszystko o niej wiesz, ale ciągle wyprowadza cię z błędu. Kiedy chce coś przemilczeć, rysuje palcem wzorki na zroszonej ściance kieliszka, żeby odwrócić Twoją uwagę, ale jeśli przyjrzyj się uważnie, dostrzeżesz melancholię. To dziewczyna o drugim dnie, które wolałaby ukryć. Mogłabym się założyć, że słucha The Smiths.

 10 A transposition of the natural relations of two elements in a proposition, for example in the sentence ‘Melissa shook her doubtful curls’ (<https://en.oxforddictionaries.com/definition/hypallage>).

'Hedonist Rose'¹¹ [topic] is a compeer full of girlish charm who would spend a whole day with you over a bottle of wine, glancing coquettishly and giggling. From the start, you think that you know everything about her but she always puts you right. When she wants to pass something over the silence she draws patterns with her finger on a dewy glass to distract your attention but if you look closer you will discern her melancholy. She is a girl with a hidden agenda which she would prefer to mask. I could make a bet that she listens to The Smiths.'

Such narrative metaphors can encompass extensive fragments of a discourse or even a whole text. In *Synamet*, such examples are tagged as narrative metaphors and only the first phrase (*Hedonist Rose is a compeer*, PERSON → PERFUME) is annotated in the ATOS's metaphor editor. We are not able to describe whole narrative metaphors using CMT, as there is no typical metaphorical transfer from the source frame onto the target frame.

6. DISCUSSION

Building a corpus of synaesthetic metaphors from blogs enabled us to observe some interesting properties of the analyzed discourse. In the analyzed texts, the ultimate "recipient" of metaphorical transfer is the SMELL frame — it can be described by lexemes taken from all other perceptual frames, e.g. *korzenno-sandalowy akord* (*spice-sandal chord*), *przejrzysty aromat* (*clear aroma*), *miękką wanilia* (*soft vanilla*), *kwaśno-gorzki zapach* (*sour-bitter smell*), *lekkie perfume* (*light perfume*). The frame that is next most likely a target in metaphorical mapping is the TASTE frame. The HEARING frame equally often serves as a target and as a source frame in synaesthetic metaphors. The frames VISION, TOUCH, and MULTIMODAL PERCEPTION are mainly source frames, while the models of verbal synaesthesia never use visual perception as the ultimate "giver" of metaphorical transfer. The widest connectivity with various target frames is typical of the VISION, TOUCH and MULTIMODAL PERCEPTION frames (in both cases—they can be mapped onto five target perceptual frames). The HEARING and TASTE frames have more limited connectivity (with only three target perceptual frames). The SMELL frame is never a source frame in verbal synaesthesia. The model of verbal synaesthesia in *Synamet* is certainly not universal. We doubt that it represents synaesthesia in general, standard Polish or in any other language, as it was created on the basis of thematic and professional blogs. Nevertheless, it clearly shows that although synaesthetic metaphors are as such universal, the schema itself is not and it can differ depending not only on various natural languages, but also on variants within one language.

11 A name of perfume by Victoria Minya.

In some aspects, the analysis of weak synaesthetic metaphors (i.e. with the non-perceptual frames as sources) in *Synamet* supports the claim made by Lakoff&Johnson (2008 /1980/) about metaphor embodiment. The universality of personification and reification clearly shows that we prefer to translate subjective sensations into something that is more common, solid and objective. At the same time, the ABSTRACT IDEAS frame, used quite frequently as a source frame for physical perception, weakens the CMT's main argument that the metaphorical transfer goes always from a more physical domain to a more abstract one.

In our opinion, pairs of frame elements support the assumptions of Petersen et al. (2007) and Sullivan (2006, 2013) that frame shifting has its own internal logic and is dependent on the structure and values of frames. For example, the complexity of an olfactory perception of perfume is represented by the complexity of musical composition (a note is an essential oil, a perfume is a composition) or by a three-dimensionality of a building. However, we think that the authors are still overlooking one important factor — the linguistic level. The synaesthetic metaphors are not just about concepts. They are also strongly determined by lexical semantics and pragmatics. The analysis of activators most frequently used in metaphorical expressions shows that some words are chosen much more often than others. For example, in *Synamet* the *nf* of the adjective *mroczny* 'obscure' is 20,813, while the *nf* of the more general synonym *ciemny* 'dark' is only 9,460 — hence, the frequency of *mroczny* is two times higher. Both words denote exactly the same feature (lack of light). At the same time, the *nf* of the adjective *ciemny* 'dark' in the National Corpus of Polish is 24,018 and *nf* of the adjective *mroczny* 'obscure' is only 5,416 — the *nf* of *ciemny* is 4 times higher. Nevertheless, metaphors using those two adjectives differ in their meaning and collocations (e.g. we can say *mroczna muzyka* 'obscure music' but *ciemna muzyka* 'dark music' sounds artificial in Polish). Therefore, our analysis of verbal synaesthesia proves that metaphor is not only a conceptual process but also involves the semantics of lexemes, which plays a very important role.

Another interesting finding is the differentiation of metaphor productivity in categories of blogs. Before we started annotating, we assumed that synaesthetic metaphors would be present in all chosen blogs, based on their subject matter. Despite having similar topics, the blogs assembled in some categories were disappointing, e.g. CUISINE (only 1% of MUs) or COFFEE (only 0.6%). On the other hand, some blogs had exceeded our expectations — e.g. blogs about perfumes were exceptionally productive (43%). We previously assumed that the wine blogs would be more plentiful in verbal synaesthesia since winespeak has a longer tradition in culture (even in Poland), and it has already developed its own jargon.

The most important finding of the analysis of synaesthetic metaphors in blogs is the abundance of atypical metaphors: mixed metaphors, entangled metaphors, and extended (narrative) metaphors. The modern theories of metaphor were constructed on the basis of recurrent, mostly artificial examples such as *coral lips*, *John is a lion* or *Juliet is the sun*. However, when we study non-prepared texts in blogs, metaphor turns out to be a more multidimensional and complex issue than the existing models assume. The theoretical models of figurative language do not consider the ways metaphors are used in discourse. Instead, they focus on abstract conceptual schemata such as *x is y* (e.g. principal subject and subsidiary subject, source domain and target domain, mental spaces and blend). As our analysis of *Synamet* shows, metaphorical expression may exceed a lexeme or a two-word phrase — it may also be an idiom, a larger passage, or even a whole text. Furthermore, when we focus on discovering the conceptual schemata of metaphors, we lose sight of the fact that the metaphorical expression has its own linguistic properties that cannot be ignored. Our studies of metaphorical expression in blogs have clearly shown the importance of a broad context in metaphor interpretation. Moreover, metaphor is often one of the elements providing textual coherence — it may organize a substantial part of the text, or even the whole of the text structure. Metaphors appear in discourse in the form of complex clusters or chains of associations (submetaphors). A single expression may also result from contamination of separate metaphors. Kövecses (2016: 3-7) claims that:

Most of the metaphors used in natural discourse are of the mixed type (...). All (target domain) concepts consist of a number of different aspects, and that each of these aspects can be conceptualized metaphorically by means of different source domains — domains that may yield incongruent images when juxtaposed.

We strongly agree with that statement. Objects of perception, such as wine, beer, music or perfumes, have many aspects, as they are combinations of various elements (e.g. a set of essential oils in perfume, or piano, bass, and drums lines combined with vocals in music), and usually stimulate several senses (e.g. color, taste, and smell of wine or beer). Nevertheless, we think that in the case of verbal synaesthesia there is yet another aspect that explains the frequent use of mixed metaphors in the analyzed blogs. Gustatory or olfactory perceptions are a subjective and complex experience. Flavor is a mixture of several sensations: not only the taste, but also the temperature, texture, and smell of the food or drink. The taste sensation is analyzed in the gustatory area located in the cortex of the brain and in the limbic system, which is responsible for emotions. Therefore, the taste sensation is extremely subjective and can evoke certain feelings and memories (cf. Skolik 2014). Mixed metaphors represent a set of subjective associations that describe the

complexity of the sensation. In our opinion, subdividing mixed metaphors into smaller pieces (as we did with *Synamet*) is only partly satisfactory because it loses some important semantic features of the metaphoric expression. For example, *Perfume is sweet and clear but it is all the same light and spatial* means something different than separate phrases: *perfume is sweet, perfume is clear, perfume is light, perfume is spatial*.

Another atypical metaphorical expression that undermines CMT is the entangled metaphor. This phenomenon was a big challenge to us because if we had considered only the semantics of this type of metaphor in annotation, its syntactic level would have been affected and the phrase would be grammatically incorrect. Entangled metaphors, as highly compressed expressions with hypallage, in the annotation process were atomized somewhat arbitrarily into separate layers. For example, in the description of the perfume *Stopniowo wycisza się cytrusowa słodycz* [*The citrus sweetness quietens down little by little*] on the semantic level we can find two metaphors: the citrus smell is sweet (TASTE → SMELL), and the sweet smell becomes quiet (HEARING → SMELL). But on the syntactic level the subject of the sentence is *słodycz* ‘sweetness’, which primarily belongs to the gustatory lexical field and on the surface, it looks like the predicate *wyciszać się* (*quieten down*) refers to the subject — the activator of the TASTE frame. The solution adopted in the *Synamet* project is not perfect as it also loses some important semantic and pragmatic features of the original entangled metaphor.

The other problematic figurative expressions are extended metaphors. The schema of extended (narrative) metaphors appears as follows: usually only one phrase is an introduction of a figurative meaning, e.g. *wine is a stripper, perfume is an opera, perfume is an old church*. In most cases, this single element in the structure is a signal that we are dealing with a metaphor. The rest of a narrative metaphor is an elaborated characterization of an object (e.g. a stripper, opera, or church) and there are no lexical elements that would violate the syntactic or semantic rules. For example, at first sight, the following fragment (4) looks like a simple description of an operatic aria:

(4) Jaka szkoda, że ta przecudnej urody aria jest tak lakoniczna, dyskretna, ulotna i delikatna — bo aż by się chciało zatracić i przepaść z kretesem w tym zniewalającym brzmieniu...

‘It’s too bad that this aria of a lovely beauty is so laconic, discreet, light, elusive and delicate — because one would like to become completely engrossed and disappear in this captivating sound.’

One could even find some synaesthetic metaphors (e.g. *aria is light*, MULTIMODAL PERCEPTION → HEARING). The problem is that the topic (referent) of this whole fragment is the perfume “Aqua di Parma”. Therefore, all epithets refer to the SMELL frame (PERFUME IS AN ARIA) and not the HEARING frame.

The extended (narrative) metaphor is the most problematic phenomenon in *Synamet*, as there is no systematic transfer from the source frame onto the target frame, e.g. (5):

(5) Jej subtelna, dyskretna zmienność przypomina leniwie zmienny układ słonecznych plamek na futerku wielkiego, żółtego kota wylegującego się na miękkim kocyku pod kwitnącą jabłónką. Kot donikąd się nie wybiera.

‘Her [Lamsa’s¹²] subtle, discreet variability resembles lazily changing layout of sunny spots on a coat of a big, yellow cat lounging on a soft blanket under a blooming apple tree. The cat is going nowhere.’

We could propose a general metaphor of PERFUME IS A CAT, but the rest of the example can hardly be interpreted using CMT. We cannot find any mappings of those sunny spots, the soft blanket or the apple tree onto the target SMELL frame.

7. CONCLUSION

In our opinion, *Synamet* proves that blogs constitute a valuable source of interesting linguistic material. They contain diversified texts with respect to style, creativity, and language proficiency. The analysis of verbal synaesthesia in blogs has brought out some important findings not only about synaesthetic metaphors but also about metaphors in general. They are not just lexeme-to-lexeme constructions or exemplifications of predicative conceptual schema *x is y*. We think that metaphor is primarily a discourse dependent phenomenon. If we do not consider the broader context, any pragmatic information, or a deep semantic analysis of lexemes, we would neglect a lot of important features of figurative language. Although there are a lot of problematic examples in the analyzed texts, we still believe that there is an internal logic to synaesthetic metaphors. We hope that further analysis of discursive examples in *Synamet* will allow us to specify the productivity of certain metaphorical schemata, as well as identify the selectional restrictions of frames, frame elements, and lexemes. We believe that the results of our analyses will permit us to establish (on a limited basis, for the time being) to what extent the metaphorization processes are systematic, and whether it is possible to construct a software system that would automatically analyze metaphors in discourse.

12 The perfume “Lamsa Arabian Oud.”

Acknowledgments

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

REFERENCES

- Buttler, Danuta 1978: *Rozwój semantyczny wyrazów polskich*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Cacciari, Cristina 2008: Crossing the senses in metaphorical language. In: Raymond W. Gibbs (ed.) 2008: *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 425–443.
- Classen, Constance 1993: *Worlds of Sense. Exploring the Senses in History and Across Cultures*. London, New York: Routledge.
- Coulson, Seana 2001: *Semantic Leaps. Frame-Shifting and Conceptual Blending in Meaning*. Cambridge: Cambridge University Press.
- Dancygier, Barbara, Eve Sweetser 2014: *Figurative Language. Cambridge Textbooks in Linguistics*. Cambridge: Cambridge University Press.
- Doroszewski, Witold 1927: Czynniki społeczny i indywidualny w rozwoju znaczeniowym wyrazów. In: *Symbolae Grammaticae in honorem Joannis Rozwadowski*, vol. 1. Kraków: Drukarnia Uniwersytetu Jagiellońskiego, 19–35.
- Fass, Dan 1991: met*: A method for discriminating. *Computational Linguistics* 17.1, 49–90.
- Fillmore, Charles 1982: Frame semantics. In: The Linguistics Society of Korea (red.) 1982: *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co., 111–137.
- Gibbs, Raymond W. 2017: *Metaphor Wars. Conceptual Metaphors in Human Life*. Cambridge: Cambridge University Press.
- Hellsten, Iina 2002: *The Politics of Metaphor: Biotechnology and Biodiversity in the Media*. Tampere: Tampere University Press.
- Jaensch, Erich. R. 1929: *Grundformen menschlichen Seins*. Berlin: O. Elsner Verlagsgesellschaft m.b.h.
- Jensen, Thomas 2017: Doing metaphor: An ecological perspective on metaphoricality in discourse. In: Beate Hampe (ed.) 2017: *Metaphor. Embodied Cognition and Discourse*. Cambridge: Cambridge University Press, 257–276.
- Judycka, Irmina 1963: Synestezja w rozwoju znaczeniowym wyrazów. *Prace Filologiczne* 18, 59–78.
- Kövecses, Zoltán 2016: A view of “mixed metaphor” within a conceptual metaphor theory framework. In: Raymond W. Gibbs (ed.) 2016: *Mixing Metaphor*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 3–15.

- Krishnakumaran, Saisuresh, Xiaojin Zhu 2007: Hunting elusive metaphors using lexical resources. In: Anna Feldman, Xiaofei Lu (eds.) 2007: *Proceedings of the Workshop on Computational Approaches to Figurative Language*. Rochester, NY: University of Rochester, 13–20.
- Lakoff, George, Mark Johnson 2008 /1980/: *Metaphors We Live By*. Chicago: University of Chicago Press.
- Mitrenga, Barbara 2014: *Zmysł smaku. Studium leksykalno-semantyczne*. Katowice: Wydawnictwo UŚ.
- Ogrodniczuk, Maciej, Katarzyna Głowińska, Mateusz Kopeć, Agata, Savary, Magdalena Zawisławska 2015: *Coreference in Polish: Annotation, Resolution and Evaluation*. Walter De Gruyter.
- Petersen, Wiebke, Jens Fleischhauer, Peter Bücker, Hakan Beşeoğlu 2008: A frame-based analysis of synaesthetic metaphors. *The Baltic International Yearbook of Cognition, Logic and Communication* 3.1, 1–22.
- Pragglejaz Group 2007: MIP: a method for identifying metaphorically used words in discourse. *Metaphor and Symbol* 22.1, 1–39.
- Ritchie, David 2017: *Metaphorical Stories in Discourse*. Cambridge: Cambridge University Press.
- Rogowska, Aleksandra 2007: *Synestezja*. Opole: Oficyna Wydawnicza.
- Semino, Elena 2008: *Metaphor in Discourse*. Cambridge: Cambridge University Press.
- Shutova, Ekaterina, Lin Sun, Anna Korhonen 2010: Metaphor identification using verb and noun clustering. In: Chu-Ren Huang, Dan Jurafsky (eds.) 2010: *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*. Beijing: Association for Computational Linguistics, 1002–1010.
- Skolik, Agnieszka 2011: *Smak w analizie sensorycznej*. Poznań: Wydawnictwo UEP.
- Steen, Gerard J., Aletta Dorst, Berenike Herrmann, Anna Kaal, Tina Krennmayr 2010: *A Method for Linguistic Metaphor Identification. From MIP to MIPVU*. Amsterdam: John Benjamins.
- Stefanowitsch, Anatol 2006: Words and their metaphors: A corpus-based approach. In: Anatol Stefanowitsch, Stefan Th. Gries (eds.) 2006: *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin: Mouton de Gruyter, 63–105.
- Sullivan, Karen 2006: Frame-based constraints on lexical choice in metaphor. *Annual Meeting of the Berkeley Linguistics Society* 32.1, 387–399.
- Sullivan, Karen 2013: *Frames and Constructions in Metaphoric Language* (vol. 14). Amsterdam: John Benjamins Publishing.
- Ullmann, Stephen 1962: *Semantics: an Introduction to the Science of Meaning*. Oxford: Blackwell.
- Werning, Markus, Jens Fleischhauer, Hakan Beşeoğlu 2006: The cognitive

- accessibility of synaesthetic metaphors. In: Ron Sun, Naomi Miyake (eds.) 2006: *Proceedings of the Twenty Eighth Annual Conference of the Cognitive Science Society*. London: Lawrence Erlbaum Associates, 2365–2370.
- Williams, Joseph M. 1976: Synaesthetic adjectives: A possible law of semantic change. *Language* 52.2, 461–478.

STRESZCZENIE

Synestezja werbalna w polskim korpusie metafor synestezyjnych

Artykuł omawia analizę metafor w blogach internetowych wykorzystanych jako materiał źródłowy w pracach nad korpusem metafor synestezyjnych (*Synamet*). Anotacja korpusu obejmowała zarówno postać gramatyczną metafor, jak i ich własności semantyczne. Podstawą metodologiczną opisu semantycznego były pojęciowa teoria metafory (CMT) Lakoffa i Johnsona (2008/1980/) oraz semantyka ramowa Fillmore'a (1982). W niniejszej pracy przedstawiamy podstawowe wyniki analiz synestezji werbalnej w korpusie, czyli model synestezji, statystyki dotyczące domen źródłowych oraz docelowych (zarówno percepcyjnych, jak i niepercepcyjnych) w metaforach synestezyjnych, a także informacje o leksemach najczęściej używanych w metaforach. Anotacja metafor synestezyjnych w korpusie ujawniła pewne słabości teorii metafory kognitywnej (Lakoff & Johnson 2008/1980/) i pokazała, że niezbędne jest stworzenie bardziej efektywnej metody opisu wyrażen metaforycznych.

Metafory synestezyjne były przedmiotem zainteresowania lingwistów, psychologów i antropologów (Ullman 1962; Williams 1976; Classens 1993). Wydaje się, że ten typ metafory jest zjawiskiem uniwersalnym dla każdego języka naturalnego. Przez typową metaforę synestezyjną rozumiemy taką metaforę, w której nazwa odnosząca się prymarnie do danego typu percepcji (np. smaku) zostaje odniesiona wtórnie do innego zmysłu (np. zapachu), por. *słodki aromat* (Buttler 1978: 115). Za Werningiem i in. (2006) definicję metafory synestezyjnej rozszerzamy o każde wyrażenie, w którym aktywowana jest domena percepcyjna (niezależnie od tego, czy jest docelowa, czy źródłowa), np. *zimny ukłon, zadziorne wino*. Zatem metafory synestezyjne, w których obie domeny są percepcyjne, określamy jako *silne metafory synestezyjne*, zaś jeśli tylko jedna domena jest percepcyjna – mówimy o *słabych metaforach synestezyjnych*.

Ponieważ w naszym przekonaniu metafora prymarnie jest zjawiskiem charakterystycznym dla dyskursu, materiałem źródłowym naszego korpusu były niepreparowane teksty z blogów poświęconych perfumom, winu, kawie, piwu, yerba mate, muzyce, masażowi, zdrowiu czy kosmetykom. Ponieważ w naszej procedurze identyfikacji metafory kluczowy był szeroki kontekst, analizowaliśmy pełne wpisy z blogów, nie tylko ich fragmenty.

Metoda anotacji przyjęta w projekcie *Synamet* wykorzystuje założenia teorii metafory pojęciowej Lakoffa i Johnsona (2008/1980/). Teoria ta mówi, że językowe wyrażenia metaforyczne jedynie odzwierciedlają pierwotne metafory pojęciowe, w których zachodzi odwzorowanie części struktury domeny źródłowej na domenę docelową, np. *MILOŚĆ TO PODRÓŻ*. Koncepcja metafory pojęciowej ma swoje wady, między innymi nie jest do końca jasne, czym są domeny pojęciowe i jaka jest ich wewnętrzna organizacja. Z tego powodu w naszych pracach zamiast domen wykorzystaliśmy pojęcie ram (Fillmore 1982). Propozycja użycia teorii ram interpretacyjnych w analizie metafor pojęciowych pochodzi od Sullivan (2006, 2013) i została podjęta przez innych badaczy, np. w pracy Dacyngier i Sweetser (2014) lub w projekcie *MetaNet: A Multilingual Metaphor Repository*. W projekcie *Synamet*

ramy zostały stworzone od podstaw. Nie było możliwe wykorzystanie angielskiego FrameNetu, ponieważ języki angielski i polski różnią się zarówno gramatycznie, jak i semantycznie. Ramy w projekcie *Synamet* zostały ściśle powiązane z badanym dyskursem – podczas analizy tekstów z blogów anotorator mógł zgłosić potrzebę dodania nowej ramy lub nowego elementu ramy już istniejącej. Zestaw ram w *Synamecie* liczy 6 ram percepcyjnych (WZROK, SŁUCH, DÓTYK, ZAPACH, SMAK i PERCEPCJA ZŁOŻONA dla wyrażeni, które na pewno opisują percepcję, ale trudno je przypisać do jednej ramy, np. gęsty, ciężki, głęboki) oraz 55 ram niepercepcyjnych (np. CZŁOWIEK, ARCHITEKTURA, ROŚLINA, PRZESTRZEŃ, PRZESTRZEŃ, CZAS, SZTUKA, SPOŁECZEŃSTWO, WOJSKO, MAGIA, HAZARD itp.). Każda element ramy jest połączony z typowym przykładem leksykalnym, np.: rama CZŁOWIEK/element EMOCJA (*gniew*), rama ZWIERZĘ/element CZĘŚĆ ZWIERZĘCIA (*pazur*).

Do anotacji metafor w korpusie *Synamet* zostało wykorzystywane narzędzie stworzone specjalnie na potrzeby projektu. Anotoratorzy korzystali podczas pracy z opracowanej dla nich instrukcji. Procedura identyfikacji metafory opierała się na podstawowych założeniach MIPVU (Steen i in. 2010), zmodyfikowanych na potrzeby naszego projektu. Anotorator najpierw musiał określić referentów danego tekstu (które nazwaliśmy tematami metaforycznymi). Było to o tyle istotne dla analizowanych tekstów, o ile często referentów można było zidentyfikować tam tylko na podstawie szerokiego kontekstu. W przykładzie *Encens Mythique d'Orient dość szybko zmierza do ambrowego akordu bazy* jednym z referentów (tematów metaforycznych) są perfumy *Encens Mysthique d'Orient*. Następnym etapem anotacji było wyodrębnienie jednostek metaforycznych (JM-ów), które odnosiły się do wyznaczonych wcześniej referentów. Jednostka metaforyczna to wyraz, fraza, zdanie, a nawet fragment tekstu, gdzie jest aktywowana przynajmniej jedna rama percepcyjna. W powyższym przykładzie mamy zatem trzy JM-y: *Encens Mythique d'Orient dość szybko zmierza* (CZŁOWIEK -> ZAPACH), *ambrowego akordu* (SŁUCH -> ZAPACH), *bazy* (ARCHITEKTURA -> ZAPACH). Anotorator następnie określał typ frazy (np. *ambrowy akord* to fraza nominalna) i wybierał ramy i ich elementy z dostępnego spisu (np. rama źródłowa SŁUCH/element ZESTAW DŹWIĘKÓW: *akord*; rama docelowa ZAPACH/element TYP ZAPACHU: *ambrowy*). Anotorator podawał także opis gramatyczny wyrazów aktywujących ramy (które określamy jako *aktywatory*) – np. rzeczownik *akord* i przymiotnik *ambrowy*.

W projekcie *Synamet* zastosowano metodę anotacji seryjnej (kaskadowej) zamiast anotacji równoległej, ponieważ eksperyment przeprowadzony podczas tworzenia Polskiego Korpusu Koreferencyjnego udowodnił, że anotacja seryjna daje lepsze rezultaty (Ogrodniczuk i in. 2015: 89). Zatem w projekcie *Synamet* teksty anotowało najpierw czterech anotoratorów, a następnie wyniki ich anotacji były sprawdzane, poprawiane i ujednolicane przez trzech superanotoratorów.

Korpus *Synamet* zawiera:

1. 684 528 słów,
2. 1414 zanotowanych tekstów z blogów,
3. 2597 tematów metaforycznych,
4. 15855 aktywatorów (wyrazów, które aktywują ramę źródłową i docelową),
5. 9217 wyrażeni metaforycznych (jednostek metaforycznych JM).

Analizowane blogi zostały podzielone na kategorie tematyczne. Ich produktywność metaforyczna znacząco się różniła (por. Tab. 1.). Test korelacji według metody Kendalla nie pokazał zależności między liczbą segmentów w danej kategorii a liczbą jednostek metaforycznych.

Jak wynika z tabeli, najbardziej produktywne są kategorie: PIWO, MUZYKA, PERFUMY, WINO.

W silnych metaforach synestezyjnych najczęściej aktywowanymi ramami źródłowymi są: WZROK, PERCEPCJA ZŁOŻONA i DÓTYK. Najbardziej typową ramą docelową jest zapach. Tabela 2 przedstawia rezydua Pearsona ($\chi^2 = 4594,4$, $df = 5$, $p < 0,001$), z których wynika, że różnice między ramami są statystycznie istotne.

Kategoria	Słowa	JM
Kawa	90 767	68
Kosmetyki	15 437	55
Kuchnia	86 966	88
Kultura	50 947	145
Masaż	80 378	41
Muzyka	127 850	2291
Perfumy	91 328	3575
Piwo	40 149	1021
Wino	54 529	935
Yerba	6 712	53
Zdrowie	40 585	49

Tabela 1. Liczba segmentów i jednostek metaforycznych w kategoriach tematycznych blogów

W Tabeli 3. prezentujemy rezydua Pearsona dla par ram w silnych metaforach synestezyjnych ($\chi^2 = 870,35, df = 25, p < 0,001$). Pary o największej frekwencji to: SŁUCH -> ZAPACH, WZROK -> SŁUCH, SMAK -> ZAPACH, PERCEPCJA ZŁOŻONA -> SMAK. Statystycznie istotne różnice są zaznaczone wytłuszczoną czcionką.

Tabela 2. Rezydua Pearsona źródłowych i docelowych ram w silnych metaforach synestezyjnych ($p < 0,001$)

	SŁUCH	PERCEPCJA ZŁOŻONA	ZAPACH	SMAK	DOTYK	WZROK
RAMA ZRÓDŁOWA	3,21714	28,31106	-54,04338	-14,28765	14,28765	32,272
RAMA DOCELOWA	-3,21714	-28,31106	54,04338	14,28765	-27,77515	-32,272

Z kolei w słabych metaforach synestezyjnych najbardziej typowe pary ram to: SPORT -> SMAK, PRZESTRZEŃ -> SŁUCH, RZECZ -> SŁUCH, POGODA -> SŁUCH, ARCHITEKTURA -> ZAPACH, ROŚLINA -> SMAK, CZŁOWIEK -> ZAPACH, por. Tabela 4 ($\chi^2 = 526,32, df = 16, p < 0,001$).

Tabela 3. Rezydua Pearsona dla par ram w silnych metaforach synestezyjnych

Źródło Cel	SŁUCH	PERC. ZŁOŻ.	ZAPACH	SMAK	DOTYK	WZROK
SŁUCH		2,644992	4,6927382	-2,25641879	1,699175	13,7583162
PERC.ZŁOŻ.	-1,471121		-0,1209030	-0,87556277	2,586414	0,8689325
ZAPACH	19,427053	-12,190764		11,30972314	-4,557809	-10,8430464
SMAK	-4,802936	9,627941	-1,4921657		2,593427	1,1327295
DOTYK	-1,759014	2,457377	-0,1445632	0,01546881		0,8377242
WZROK	-3,313865	7,055469	2,2409309	-1,60628493	3,542616	

Ramy docelowe różnią się zestawem ram źródłowych, z którymi się łączą – najbardziej zróżnicowany zestaw ram źródłowych charakteryzuje słabe metafory synestezyjne, w których ramą docelową jest ZAPACH (44 ram). Rama SŁUCH łączy się z 37 niepercepcyjnymi ramami źródłowymi, SMAK – z 27, zaś WZROK tylko z 16. Ramy DOTYK oraz PERCEPCJA ZŁOŻONA nigdy nie występują jako ramy docelowe w słabych metaforach synestezyjnych.

Tabela 4. Rezydua Pearsona dla par ram w słabych metaforach synestezyjnych

Cel	Źródło	SŁUCH	SMAK	ZAPACH
ARCHITEKURA		-2,818659	-4,524547	6,266083
CZŁOWIEK		-5,1182337	0,1791543	4,6642333
POGODA		6,544754	-4,912604	-2,218539
PRZESTRZEŃ		8,063878	-4,223139	-4,196844
ROŚLINA		-4,640675	6,251759	-0,640944
RZECZ		7,229340	-2,326573	-4,929763
SPORT		-6,691135	16,449416	-6,870617
SZTUKA		-0,4312885	-2,1051118	2,0886845
UBRANIE		1,315318	-1,351310	-0,154747

Silne metafory synestezyjne różnią się zasadniczo zestawem elementów ram źródłowych. Najwięcej elementów jest aktywowanych w ramie WZROK – 52, w ramie SŁUCH – 39 elementów, w ramie DOTYK – 31 elementów, w ramie SMAK – 21 elementów. Najmniej elementów jest wybieranych z rami PERCEPCJA ZŁOŻONA – tylko 18.

Aktywatory (czyli wyrazy, które aktywują rami) w narzędziu do anotacji w projekcie *Synamet* mają dwie formy – tekstową oraz formę podstawową (słownikową). Dzięki takiemu zabiegowi powstał mini słownik aktywatorów metafor synestezyjnych. Zbiory aktywatorów powiązanych z sześcioma ramami percepcyjnymi różnią się zarówno ogólną liczbą, jak i zasobnością klas gramatycznych. Najwięcej aktywatorów odnosi się do ram ZAPACH (458) oraz SŁUCH (426). Najmniej wyrazów aktywuje ramę PERCEPCJA ZŁOŻONA (142), por. Tabela 6.

Tabela 6. Aktywatory ram percepcyjnych w Synamecie

RAMA	Adj	Adv	N	V	LICZBA LEKSEMÓW
ZAPACH	110	3	336	9	458
SŁUCH	86	17	289	34	426
WZROK	158	23	122	84	387
SMAK	98	13	210	9	330
DOTYK	88	13	35	44	180
PERCEPCJA ZŁOŻONA	67	19	35	21	142

W analizowanych metaforach synestezyjnych aktywatory o najwyższej frekwencji to:

- 1) w ramie SŁUCH: *nuta* (nf= 304,005), *kompozycja* (nf= 111,636), *akord* (nf= 67,486), *utwór* (nf= 63,702), *brzmienie* (nf= 58,656);
- 2) w ramie PERCEPCJA ZŁOŻONA: *lekki* (nf= 144,433), *ciężki* (nf= 51,718), *lekko* (nf= 44,15), *głęboki* (nf= 35,320), *głębka* (nf= 17,029);
- 3) w ramie ZAPACH: *zapach* (nf= 247,240), *aromat* (nf= 61,810), *perfumy* (nf= 61,179), *zapachowy* (nf= 29,012), *róża* (nf= 21,120);
- 4) w ramie SMAK: *wino* (nf= 116,682), *słodki* (nf= 71,270), *słodycz* (nf= 48,465), *smak* (nf= 48,565), *gorczyzka* (nf= 37,842);

5) w ramie DOTYK: *delikatny* (nf= 66,855), *ciepły* (nf= 47,303), *ostrzy* (nf= 42,888), *miękki* (nf= 20,182), *gładki* (nf=19,552);

6) w ramie WZROK: *tło* (nf= 54,872), *wysoki* (nf= 44,150), *czysty* (nf= 35,950), *barwa* (nf= 25,859), *pojawiać się* (nf= 25,859);

Narzędzie stworzone do anotacji wyrażeń metaforycznych w projekcie *Synamet* zostało dostosowane do teorii metafory pojęciowej, to znaczy do opisu wyrażenia metaforycznego w postaci typowego dualnego schematu X TO Y, np. *Womanity (perfumy) gaśnie powoli* – zapach to blask, *Tyle innych burgundów cierpi na sklerozę i haluksy* – WINO TO CZŁOWIEK. Jednakże podczas anotacji tekstów z blogów wyróżniliśmy trzy typy wyrażeń metaforycznych, których nie dało się odwzorować w postaci takiego schematu: metafory wieloramowe (por. Kövecses 2016: 3), metafory szkatułkowe i metafory narracyjne (Gibbs 2017: 51; Ritchie 2017).

Metafory wieloramowe są typowym zjawiskiem w tekstach składających się na korpus *Synamet*. Jest to zjawisko polegające na tym, że jeden temat metaforyczny (referent) jest opisywany za pomocą wyrażeń aktywujących kilka różnych ram źródłowych, por. przykład 1.

(1) Santo Stefano i Santo Stefano Riserva to wina głębokie, długie, eleganckie, o twardej ramie garbników, a zarazem mięsistej, sutej materii.

Anotacja tego typu metafor wymagała rozbicia ich na mniejsze jednostki, w których powtarzał się temat metaforyczny, natomiast inne były ramy źródłowe, np. *wina głębokie* [PERCEPCJA ZŁOŻONA -> SMAK], *wina długie* [WZROK -> SMAK], *wina eleganckie* [UBRANIE -> SMAK].

Metafory szkatułkowe to bardziej skomplikowany typ, który polega na tym, że za stosunkowo prostą powierzchniowo frazą kryje się kilka metafor i metonimii, a także mamy do czynienia z hypallagą, ponieważ relacje składniowe nie odzwierciedlają faktycznych relacji semantycznych. Na przykład w wypowiedzeniu *Ta płyta to mroczna podróż przez 20 utworów* przymiotnik *mroczny*, pozornie odnoszący się do rzeczownika *płyta*, tak naprawdę opisuje metaforycznie wrażenia słuchacza muzyki znajdującej się na płycie. Druga metafora to zobrazowanie słuchania muzyki jako podróży. Mamy tu też metonimię – *płyta* za muzykę i za słuchanie muzyki. Anotacja tego typu wyrażeń metaforycznych polegała na rozbiciu ich na trzy frazy: *płyta to podróż*, *mroczna podróż*, *płyta (...) mroczna*.

Metafory narracyjne to bardzo rozbudowane metafory, wychodzące poza granice zdania, przedstawiające temat metaforyczny w postaci opowieści (Ritchie 2017: 242), np.:

Hedonist Rose to pełna dziewczęcego uroku towarzyszka, która spędzi z Tobą przy butelce wina cały dzień, od niechcenia rzucając zalotne spojrzenia i chichocząc. Od początku myślisz, że wszystko o niej wiesz, ale ciągle wyprowadza cię z błędu. Kiedy chce coś przemilczeć, rysuje palcem wzorki na zroszonej ścianie kieliszka, żeby odwrócić Twoją uwagę, ale jeśli przyjrzyj się uważnie, dostrzeżesz melancholię. To dziewczyna o drugim dnie, które wolałyby ukryć. Mogłabym się założyć, że słucha The Smiths.

W edytorze metafor w projekcie *Synamet* jedynie fragment takiej metafory dawał się zanotować (np. *Hedonist Rose to (...) towarzyszka* – CZŁOWIEK -> ZAPACH). Poza tym nie dawało się opracować w ten sposób całej metafory narracyjnej, ponieważ nie było wyraźnego schematu odwzorowania elementów narracji z ramy źródłowej w ramie docelowej.

Tworzenie korpusu metafor synestezyjnych pozwoliło zaobserwować wiele interesujących cech analizowanych tekstów. Przede wszystkim, totalnym „biorcą” – najbardziej typową ramą docelową okazała się rama ZAPACHU, którą opisywały wyrażenia należące prymarnie do wszystkich pozostałych ram percepcyjnych. Ramą docelową była także stosunkowo często rama SMAKU. Natomiast ramami źródłowymi o najwyższej frekwencji były: WZROK, DOTYK ORAZ PERCEPCJA ZŁOŻONA.

Opis połączeń par ram percepcyjnych pozwoliło opracować model synestezji językowej w korpusie *Synamet*. W odróżnieniu od badaczy, którzy wcześniej podejmowali próby stworzenia modelu

synestezji, nie uważamy naszego modelu za uniwersalny, co więcej – prawdopodobnie uległby on zmianie nawet w obrębie polszczyzny, gdyby korpus składał się z innego typu tekstów. Zatem, choć sama metafora synestezyjna może być uważana za zjawisko uniwersalne, model synestezji zależy już od danego języka i jego odmiany.

Analiza słabych metafor synestezyjnych częściowo potwierdza tezę Lakoffa i Johnsona o metaforze ucieleśnionej – wysoka frekwencja personifikacji i reifikacji świadczy, że preferujemy opisywanie bardzo subiektywnych odczuć smakowych lub olfaktorycznych jako konkretnych, trwałych obiektów lub jako ludzi. Jednak metafory synestezyjne, w których jako rama źródłowa wykorzystywane są pojęcia abstrakcyjne, osłabiają tę tezę.

Połączenia ram interpretacyjnych i ich elementów w korpusie *Synamet* potwierdzają spostrzeżenie Petersena i in. (2007) oraz Sullivan (2006, 2013), że transfer metaforyczny nie jest przypadkowy, ale cechuje go pewna logika. Na przykład złożoność percepcji olfaktorycznej aromatu perfum jest obrazowana za pomocą kompozycji muzycznej (nuta to olejek eteryczny, perfumy to kompozycja) lub jako trójwymiarowa konstrukcja. Wydaje nam się jednak, że skupienie się na poziomie pojęciowym sprawia, że gubi się element językowy. Wyrażenia metaforyczne w badanych tekstach są zdeterminowane także przez czynniki semantyczne i pragmatyczne. Na przykład w korpusie *Synamet* przymiotnik *mroczny* ma dużo wyższą frekwencję niż przymiotnik *ciemny* (odwrotnie niż w NKJP). Chociaż oba przymiotniki aktywują ten sam element ramy wzrok (*BRAK ŚWIATŁA*), to jednak różnią się nacechowaniem i łączliwością.

Kolejnym ważnym wynikiem przeprowadzonych badań jest stwierdzenie wyraźnej różnicy w produktywności metaforycznej anotowanych blogów. Niektóre kategorie mają wyjątkowo niską liczbę metafor synestezyjnych (np. *KUCHNIA* lub *KAWA*), z kolei inne cechuje niezwykle wysoka produktywność (*PERFUMY*, *WINO*, *PIWO*, *MUZYKA*).

Najciekawszym wynikiem analiz jest wyróżnienie nietypowych wyrażeń metaforycznych (z punkty widzenia teorii metafory pojęciowej). Jak wynika z anotacji przeprowadzonej w korpusie *Synamet*, metafory tekstowe to nie tylko związek dwóch wyrazów, lecz także rozbudowane frazy, a zdarza się, że przekraczają one granice zdania i mogą być bardzo obszernymi narracjami obejmującymi duże fragmenty tekstu lub nawet cały tekst. Postać tekstowa metafor bywa w związku z tym niezwykle skomplikowana i wymaga wieloetapowej anotacji. Badanie niepreparowanych tekstów z blogów pokazało ponadto, jak ważne jest uwzględnienie bardzo szerokiego kontekstu w procesie identyfikacji metafory.

BIBLIOGRAFIA

- Buttler, Danuta 1978: *Rozwój semantyczny wyrazów polskich*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Classens, Constance 1993: *Worlds of Sense. Exploring the Senses in History and Across Cultures*. London, New York: Routledge.
- Dancygier, Barbara, Eve Sweetser 2014: *Figurative Language. Cambridge Textbooks in Linguistics*. Cambridge: Cambridge University Press.
- Fillmore, Charles 1982: Frame semantics. W: The Linguistics Society of Korea (red.) 1982: *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co., 111–137.
- Gibbs, Raymond W. 2017: *Metaphor Wars. Conceptual Metaphors in Human Life*. Cambridge: Cambridge University Press.
- Kövecses, Zoltán 2016: A view of “mixed metaphor” within a conceptual metaphor theory framework. W: Raymond W. Gibbs (red.) 2016: *Mixing Metaphor*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 3–15.

- Lakoff, George, Mark Johnson 2008 /1980/: *Metaphors We Live By*. Chicago: University of Chicago Press.
- Ogrodniczuk, Maciej, Katarzyna Głowińska, Mateusz Kopeć, Agata Savary, Magdalena Zawislawska 2015: *Coreference in Polish: Annotation, Resolution and Evaluation*. Berlin: Walter De Gruyter.
- Petersen, Wiebke, Jens Fleischhauer, Peter Bücken, Hakan Beseoglu 2008: A Frame-based analysis of synaesthetic metaphors. *The Baltic International Yearbook of Cognition, Logic and Communication* 3.1, 1–22.
- Ritchie, David 2017: *Metaphorical Stories in Discourse*. Cambridge: Cambridge University Press.
- Steen, Gerard J., Aletta Dorst, Berenike Herrmann, Anna Kaal, Tina Krennmayr 2010: *A Method for Linguistic Metaphor Identification. From MIP to MIPVU*. Amsterdam: John Benjamins.
- Sullivan, Karen 2006: Frame-based constraints on lexical choice in metaphor. *Annual Meeting of the Berkeley Linguistics Society* 32.1, 387–399.
- Ullman, Stephen 1962: *Semantics: an Introduction to the Science of Meaning*. Oxford: Blackwell.
- Wering, Marcus, Jens Fleischhauer, Hakan Beseoglu 2006: The cognitive accessibility of synaesthetic metaphors. W: Ron Sun, Naomi Miyake (red.) 2006: *Proceedings of the Twenty Eighth Annual Conference of the Cognitive Science Society*. London: Lawrence Erlbaum Associates, 2365–2370.
- Williams, Joseph M. 1976: Synesthetic adjectives: A possible law of semantic change. *Language* 52.2, 461–478.