MAGDALENA ANNA ZAWISŁAWSKA

University of Warsaw, Poland
zawisla@uw.edu.pl

SYNAMET — A MICROCORPUS OF SYNESTHETIC METAPHORS. PRELIMINARY PREMISES OF THE DESCRIPTION OF METAPHOR IN DISCOURSE

Abstract

This article describes the preliminary premises of metaphor annotation in SYNAMET — the developing microcorpus of synesthetic metaphors. The analysis is based on the CLST theory (Context-Limited Simulation Theory) put forward by D. Ritchie. According to this theory, the metaphor's vehicle may activate various types of associations between words: semantic relations, perceptual sensations, or emotional simulations. The range of potential associations evoked by the vehicle is limited by the topic, i.e. the lexical context in which the metaphor appears. The relations between the vehicle and the topic may be presented in the form of a semantic frame.

To reconstruct the frames within the project, linguistic works devoted to sensory perception-vision, hearing, smell and taste- will be utilized. The corpus annotation will consist of the following stages:

1) metaphor identification
2) indication of the metaphor cluster (CM) — a phrase or a passage of the text, centered around one referent
3) isolation of the metaphorical units (MU) — word forms or phrases combining lexemes primarily belonging to different perceptual frames.

The outcome of the MU analysis will include: a general metaphorical scheme of the MU, lexical items activating the frame of the MU (together with their grammatical description), a detailed metaphor scheme of the MU, and the semantic and grammatical categorization of the MU.

Keywords: metaphor; corpus; synesthesia; frames

1 Metaphor and Corpus

As a phenomenon which involves both conceptual systems and language, metaphor has been of interest to researchers from various fields, e.g. psychology, neurology, literary studies, linguistics, natural language processing, etc. Consequently, there is now a substantial body of scientific literature devoted to this topic. There has recently been a growing interest in corpus-based studies of metaphor. The following works may constitute practical guides for researchers investigating metaphor in discourse: *Researching and Applying Metaphor* (Cameron, Low 1999), *Discourse approaches to metaphor research* (Deignan 2005), *Corpus-Based Approaches to Metaphor and Metonymy* (Stefanowitsch, Gries 2008), Elena Semino's *Metaphor in Discourse* (2008), *Metaphor Analysis. Research Practice in Applied Linguistics, Social Sciences and the Humanities* (Cameron,

Maslen 2010). In addition to these publications, metaphors in various discourse types have been comprehensively studied, e.g. in ideological discourse (Goatly 2006), in political and economic discourse (Skorczynska, Deignan 2006), in advertising (Lundmark 2005), and in educational discourse (Cameron 2003).

With regards to the Polish language, there is still only a limited number of studies of metaphor in (larger or smaller) corpora. P. Wróblewski's book *Struktura, typologia i frekwencja polskich metaphor* (1998), in which the author studies the frequency of metaphors in four 100,000-words corpora, is definitely noteworthy. Another significant monograph is *Metaforyka Leśmiana. Analiza lingwistyczna* (2011) by W. Cockiewicz, which provides a thematically-organized dictionary of the metaphors occurring in B. Leśmian's poetic texts. Moreover, E. Badyda, in her dissertation *"Upadły anioł zmysłów"? Metaforyka zapachu i percepcji węchowej we współczesnej polszczyźnie* (2013), has examined examples excerpted from the PWN corpus and the IPI PAN corpus, as well as from texts from the Internet.

Corpus-based research of metaphor provides insight into the textual functions of metaphors. Theoretical works on metaphor tend to regard it as a syntactic scheme X is Y. J. Soskice (1985, p. 19) argues that many disputes over metaphor stem from the fact that researchers rarely specify whether they are discussing the syntactic form of a metaphor, or its logical structure. In addition to this, they often confuse the two.

Soskice claims that taking into account solely the X is Y structure of metaphorical expressions implies that, grammatically speaking, metaphor is always an assertion and its vehicle is always a predicate. In contrast, Soskice believes that metaphor does not manifest itself in just one unique form, since it is identified not only on the basis of syntactic criteria, but also on semantic and pragmatic ones. The multiformity of metaphors in texts is also noted by A. Goatly and L. Cameron.

It follows clearly from Goatly's (1997) analysis of examples from literature and from the *Bank of English* that metaphors are expressed not only by nouns, but by other parts of speech as well. Cameron (2003, p. 75), in her corpus-based study, has distinguished metaphors of different grammatical forms (verb and preposition metaphors), and comparison metaphors involving two incongruous domains arranged to form an explicit comparison. Furthermore, Cameron (2003) and Deignan (2005) have observed that metaphorical expressions displaying the X is Y structure are actually quite rare in texts. Cameron (2003, p. 88–89) has compared the number of metaphors expressed by different linguistic forms, and it has enabled her to determine that almost 50% of metaphors come in the form of verbs. Metaphorical expressions involving nouns (of the type A=B, e.g. *Man is a wolf*) make up no more than 5% of the corpus.

Deignan (2005, p. 178) has restricted her research on the ways metaphors are expressed to lexemes involving the domain of plants and, more precisely, she focused on the English word *blossom*, which may either be a collective noun denoting flowers, or a verb meaning 'to produce flower(s).' As she reports, the noun had 167 occurrences in the literal meaning and was used only twice in the figurative sense, whereas the verb occurred 5 times in its literal meaning, and as many as 55 times in the figurative meaning.

As P. Wróblewski (1998, p. 31) rightly points out, in the description of metaphorical expressions the conceptions that prove to be the most useful are those that account for the lexicalized elements of metaphors. However, if we were to consider any hypothetical elements of the metaphor (e.g. as happens in the case of an obligatory reduction of all metaphorical expressions to the X is Y schema), the study's accuracy would diminish, since it would in fact consist of a subjective and arbitrary interpretation of metaphor.

It is insufficient to analyze metaphor only in the existing corpora, since only a few selected lexemes may be examined in such a way. Corpus search engines are practically useless for metaphor researchers: they cannot assist in the examination of the complexity of metaphor, neither can they provide information on the grammatical properties of metaphor and on the productivity of metaphorical schemata. A concordance will not help to single out innovative metaphors that may structure complex narration-access to the whole discourse is indispensable in order to do that.

Therefore, a separate corpus of metaphors is a better tool for metaphor analysis.

In the only available corpus of metaphors — *VU Amsterdam Metaphor Corpus* (`http://www2.let.vu.nl/oz/metaphorlab/metcor/search/`) — attention is fixed exclusively on analyzing lexemes that in a given context acquire either a figurative or a literal meaning. The corpus employs the *BNC Baby* (a 4-million subcorpus of the British National Corpus). Among the tagged texts one may find academic texts, news texts, fiction, and conversations. The authors divided the metaphors into the following categories:

1) *clear metaphors*: *direct* (displaying "a contrast between the contextual and a more basic meaning"), *indirect* ("a contrast as well as comparison between the contextual and a more basic meaning"), and *implicit metaphors* ("due to an underlying cohesive grammatical and/or semantic link in the discourse which points to recoverable metaphorical material")
2) *personifications*
3) *borderline metaphors*. This last category includes expressions whose metaphorical status was unclear to the annotators.

Additionally, metaphor signals (such as e.g. *like, as, as if, so-called*) have been taken into consideration. The search engine available on the project's website so far allows queries concerning just one category — personification. The other search modules allow one to look for metaphor-related words, or to display concordances in which the keyword appears both in literal and figurative meanings. Cf. Tab. 1.

Table 1:

loc	context
conv	You crafty **cow** (figurative meaning)
fic	Almost every man is accompanied by a goat or a **cow** held on a short leather thong or controlled by a thin stick
fic	Around them graceful Masai women in long leather skirts are sweeping the boma and plastering the huts with **cow** dung
news	Small children can sometimes be persuaded to eat **cow** cake in the same way that old people will eat cat food but the chief result of all this farming was to
news	Nowadays we all swoon with pleasure at the sight of a **cow**

MetaNet: *A Multilingual Metaphor Repository* (`https://metanet.icsi.berkeley.edu/metanet/`) is the second interesting project, though it is still in its preliminary stages of preparation. As a result of the project's guidelines, the work is aimed at designing a system that will be able to extract metaphors from texts, and to interpret them automatically. The project involves four languages: English, Persian, Russian, and Spanish, and its authors (S. Narayan, B. Bergen, C. Baker, J. Feldman, G. Lakoff, E. Shutova, L. Azis-Zadeh, L. Boroditsky, T. Matlock, E. Sweetser) employ the conceptual metaphor theory, formulated by G. Lakoff and M. Johnson, as well as the FrameNet ontology. Therefore, it is not intended to be a corpus of metaphors, but rather an item software dedicated to analyzing metaphors in texts.

Corpora of metaphors designed for the English language are undoubtedly interesting resources but they cannot introduce one to the unique features of metaphors in Polish. Therefore, it is necessary to build a separate corpus of metaphors dedicated to the Polish language that would help to create a new metaphor classification, according to semantic and grammatical properties, and to propose a more formalized method of metaphor description. These are the main goals of the project SYNAMET — A *Microcorpus of Synesthetic Metaphors. Towards a Formal Description and Efficient Methods of the Analysis of Metaphors in Discourse.*[1]

Magdalena Anna Zawisławska   - 110 -

SYNAMET — A Microcorpus of Synesthetic Metaphors. Preliminary Premises of the Description...

2   Synesthetic Metaphor

It would seem that research on such a complex phenomenon as metaphor should start from a restricted research area, and focus on analyzing authentic texts instead of fabricated ones. It appears that synesthetic metaphor may prove to be a valuable material for preliminary research.

D. Butler (1978, p. 115) defines synesthetic metaphor as "using one name for various types of sensory perception." In psychology, *synesthesia* pertains to "a co-occurrence of sensations coming from various pathways while stimulating one pathway, e.g. hearing given tones when looking at given colors, and vice versa" (Rogowska 2007, p. 18). Synesthesia in the Polish language has been studied, from a diachronic point of view, by Irmina Judycka (1963). Other authors have also undertaken fragmentary studies of this topic, for example, B. Mitrenga's PhD thesis *Zmysł smaku w polskiej leksyce i frazeologii* or E. Badyda's monograph *„Upadły anioł zmysłów"? Metaforyka zapachu i percepcji węchowej we współczesnej polszczyźnie* (2013). Synesthetic metaphor in various discourse types has also been the subject matter of several papers, e.g. K. Termińska (1992), M. Witucka (1998), A. Rosińska (2007), E. Biłas-Pleszak (2007), A. Najdecka (2013), A. Prochowicz (2013).

There are several reasons for investigating synesthetic metaphor. It has been studied not only by linguists, but also by psychologists and anthropologists. It appears to be universal for all natural languages. S. Ullman (1962, p. 861) proposed the following model of synesthesia: *smell / taste – hearing / vision – touch*, which means that we usually speak of gustatory and olfactory sensations in terms of hearing and seeing, while auditory and visual sensations are conceived of in terms denoting tactile experiences. C. Classens (1993) likewise elaborated a (supposedly universal) schema of synesthetic metaphors: *hearing – vision – smell – taste – touch*. It follows from this model that auditory sensations are, as a rule, expressed in terms of visual perception, visual ones-in terms of olfactory perception, olfactory ones-in terms of gustatory perception, and gustatory ones-in terms of tactile perception. However, studies carried out by Wering, Fleischhauer and Beşeoğlu (2006) have plainly shown that the models put forward by Ullman and Classens are by no means universal, and that they need to be constructed separately for each language. The benefits of choosing synesthetic metaphor as an object of study involve both a precise indication of the scope of research, and a guarantee of rich and diversified material for analysis. In some approaches, synesthetic metaphor embraces all the expressions containing any word primarily denoting a perceptual domain, i.e. either phrases such as *jasny dźwięk* ('bright sound'), *słodki zapach* ('sweet smell'), or constructions such as *jasny umysł* ('clear mind'), *ciemna sprawa* ('shady business').

The synesthetic metaphor category embraces various subtypes. Judycka cites the taxonomy proposed by E. R. Jaensch, who distinguished synesthesia in the narrow sense (expressions denoting only the perceptual sphere) and synesthesia in the broad sense (emotional synesthesia, such as e.g. *czarna rozpacz* 'dark despair,' conceptual symbols, and the so-called complex synesthesia). Judycka herself discriminates between word synesthesia (i.e. etymological, entrenched), e.g. *lęk* 'fear' (Proto-Slavic * *lęk-ǫ* 'I bend'), and phraseological synesthesia. She also makes a distinction between simple and complex synesthesia, the latter category consists of merging the sensations coming from different perceptual domains, e.g. *chudy* 'thin/lean,' *tłusty* 'fat/greasy' (por. Judycka 1963, p. 59–60). 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 deal with *a strong synesthetic metaphor* (Werning, Fleischhauer, Beşeoğlu 2006).

In order to collect the most varied material possible, the project will employ the broad sense of the term *synesthetic metaphor*. Hence, the following types of synesthetic metaphors shall be included, cf.

1. Simple strong synesthesia, e.g. *jasny dźwięk* 'clear tone', *ostry zapach* 'sharp smell.'
2. Simple weak synesthesia, e.g. *czarna melancholia* 'dark melancholy', *słodkie życie* 'sweet life.'

3. Complex strong synesthesia, e.g. *ciężki zapach* 'heavy odor', *lekki smak* 'light taste.'
4. Complex weak synesthesia, e.g. *tłusty dowcip* 'dirty joke', *chudy intelekt* 'meager wit.'

However, *the word synesthesia* (as Judycka calls it) will be excluded, since expressions of this type require etymological analysis. Since metaphor basically works in a context, the project analyzes whole texts, and not excerpts. Nonetheless, owing to copyright restrictions, the complete corpus will allow access to only the fragments containing metaphors.

Since the main topic is synesthetic metaphors, the project concentrates on such texts where the probability of them occuring is the highest. In principle, it excludes the possibility of employing the National Corpus of Polish. Instead, texts excerpted from blogs devoted to perfume (SMELL), wine, beer, cigars, Yerba Mate, tea, or coffee (TASTE, SMELL, VISION), as well as culinary blogs (TASTE, VISION), music blogs (HEARING), art blogs (VISION), massage and wellness blogs (TOUCH) are used. Without doubt, the corpus will not be fully balanced, but at this stage this is not a set objective, since microcorpora are essentially thematically organized.

3 The Theoretical Basis of Metaphor Description in Discourse

There is a considerable number of theories of metaphor — the classical substitution and simile theory, the reinterpretation theory (Searle 1993), the interaction theory (Richards 1936), the theory of metaphor as predication (Bogusławski 1971, Dobrzyńska 1984, 1994, Wierzbicka 1971, Arutjunowa 1981, Sedivy 1997), the theory of metaphor as categorization (Glucksberg, Keysar 1993), the conceptual metaphor theory (Lakoff, Johnson 1988), and the conceptual integration theory (Fauconnier, Turner 2002).

The problem that the majority of these proposals (the older ones in particular) pose is that they are usually highly theoretical, and either almost deprived of illustrative material, or else based on a few lexicalized metaphorical expressions. As a consequence, researchers study metaphors which are taken out of context or set in a very limited context, whereas studies of metaphorical expression in discourse have clearly shown the importance of context in metaphor interpretation. Moreover, metaphor is often one of the elements providing textual coherence — it may organize a substantial body of the text, or even the whole of the text structure. It is virtually impossible to apply simple metaphorical schemata (such as X is Y) put forward by many researchers, since metaphors appear in discourse in the form of complex clusters or chains of associations (submetaphors). A single expression may also result from the contamination of separate metaphors. Nor is the difference between metaphor and metonymy clear-cut (cf. Semino 2008).

Many Polish studies on metaphor use the interaction theory of metaphor. The problematic issue with this theory is that it lacks a clear definition of *interaction*. I. Richards (1936) employs the terms *interanimation* and *cooperation*, while M. Black (1993) speaks of *isomorphism*. In T. Dobrzyńska's interactive conception of metaphor, metaphor comes down to predication (cf. Dobrzyńska 1994).

The majority of recent corpus-based studies utilize either the conceptual metaphor theory, formulated by G. Lakoff and M. Johnson (1988), or the conceptual integration theory (Fauconnier, Turner 2002). Lakoff and Johnson view metaphor as a primarily conceptual phenomenon which consists of mapping across domains (from the source domain onto the target domain). However, it has not been precisely stated what the term *conceptual domain* represents, and how the domain structure is to be reconstructed. Moreover, the metaphors are presented as a very general schemata X is Y, which results in the metaphors' grammatical properties exhibited in the text being ignored. On the other hand, the conceptual integration theory formulated by G. Fauconnier and Mark Turner is highly intricate, and it is quite difficult to put it into practice. The definition of mental space is as vague as that of the conceptual domain. Moreover, the analysis of metaphor as a linguistic phenomenon is, beyond doubt, insufficient.

Slightly more precision is offered by the CLST (*Context-Limited Simulation Theory*) put forward by D. Ritchie (2006, 2011). This theory is based on D. Sperber and D. Wilson's *Relevance Theory*, as well as on the *Perceptual Simulation Theory* proposed by L. Barsalou (cf. Ritchie 2006). The CLST admits the essential role that the context plays in metaphor analysis — the participants of a communication event create and sustain the discourse world. The metaphor's *vehicle* may activate various associations: semantic relations between certain words, sensations, or emotional simulations. The range of potential associations evoked by the vehicle is limited by the *topic*, i.e. the lexical context in which the vehicle appears. The relations between the vehicle and the topic may be presented in the form of a semantic frame.

Frames, which constitute mental structures organizing human experiences, are of interest to many scientific fields, cf. e.g. R. Schank and R. Abelson (1977), as well as M. Minsky (1975) for artificial intelligence (AI), as well as G. Bateson for anthropology, E. Goffman for sociology (as cited in Tannen 1993, p. 15). It is worth mentioning that the term *frame* is understood in various ways by various researchers. Also, different terms are sometimes used, such as: *schema*, *script*, *frame*, *semantic frame*, *cultural model*, *cognitive model*, *domain*, *gestalt* (cf. Kövecses 2011, p. 105, Tannen 1993, p. 15). However, the essential point is that a frame is seen as an ordered structure within which there are categories (*slots*) and their values (*fillers*). Frames are hierarchically organized. They may represent various events, states, objects, and situations. Frames are either static or dynamic (cf. Schank and Abelson scripts).

Since we are interested exclusively in the linguistic embodiments of frames, we need to evoke Ch. Fillmore's definition of a frame. Fillmore posits that the meaning of lexical units, phrases, and grammatical and syntactic constructions resides in schematic phenomena, such as our beliefs, experiences or typical actions (Fillmore 1982). Thus, the metaphorization process may in this approach be described as *frame shifting*. S. Coulson (2001) defines this phenomenon as a „semantic reanalysis process that reorganizes existing information into a new frame." This means that some elements of a frame evoking specific sensations (e.g. smell) as its topic may become reorganized under the influence of a vehicle activating a frame of some other sensory perception (e.g. hearing). Cf. *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 example, the lexical units *smell*, *chypre* and *musk* are a topic that activates the SMELL frame, while phrases *breath* and *purrs* activate the HEARING fame.

4   The Problems of Metaphor Identification in Discourse

The principle challenge to be faced is the creation of a metaphor identification procedure. The essential problem lies with distinguishing literal and non-literal usages. The metaphoricity of certain expressions may be perceived quite differently by various annotators. Hence, metaphor forms a continuum ranging from the most typical phenomena, through more peripheral ones, up to the borderline cases residing on the verge of the literal and the figurative meanings.

Metaphor researchers usually postulate two separate categories: novel, alive, creative metaphors, as opposed to dead, lexicalized, entrenched metaphors — set expressions or lexemes whose figurative meanings have already been registered in dictionaries. Black (1993) argues that only the former should be focused on. By contrast, Lakoff and Johnson (1988) pay attention to lexicalized metaphors, since, in the authors' view, it is precisely these metaphors — used repeatedly and automatically in everyday discourse — that disclose the metaphorical nature of our conceptual system.

Discriminating lexicalized and novel metaphors in discourse is by no means an easy task. Text analysis has shown that in some contexts a seemingly dead metaphor may become alive. S. Knudsen (2003) has described it as the opening of a closed metaphor. C. Müller (2008) has proposed disposing of this division into the two classes. Instead, she has distinguished dead

metaphors (where conceiving the target domain in terms of the source domain is entirely inactive in an average language speaker), sleeping metaphors (with low-level activation), and alive metaphors (high activation level). This discussion highlights the fact that it is difficult to determine the borders of metaphor in a text.

In our project, we employ the proposal put forward by L. Cameron and R. Maslen (2010), who introduced the term *potential metaphor*. In this approach, an expression does not need to be seen as metaphorical by all its users, neither is it assumed that speakers create metaphors consciously. Adopting this solution will enable the inclusion of both alive and dead metaphors in the project.

There are several different procedures for the recognition of metaphors in discourse. For instance, we have at our disposal the *met* system* (Fass 1991) or Krishnakumaran and 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 metaphor recognition procedure elaborated by Ekaterina Shutova, Lin Sun and Anna Korhonen (2010) draws on the systematic character of conceptual metaphor as it is seen by George Lakoff and Mark Johnson (1988). In this method, the starting point consists of determining the basic set of metaphorical expressions that represent the source domain to target domain mappings. Subsequently, nouns denoting various concepts belonging to the target domain, and verbs denoting situations evoking the source domain are collected. In this way, a verbal lexicon pertaining to the source domain comes into being. The final stage involves browsing the corpus in order to note the expressions from the target domain with verbs from the source domain. However, this procedure is far from idea, since it fundamentally leads to equating every metaphorical expression to a noun-verb combination.

The most widely-known metaphor recognition procedure is MIPVU (Steen *et al.* 2010), which constitutes a modified and elaborated version of the MIP (*Metaphor Identification Procedure*) as proposed by the Pragglejaz group (Pragglejaz 2007, Semino 2008, p. 11–12). The name of the group is an acronym formed from the first names of the its members, i.e. **P**eter Crisp, **R**ay Gibbs, **A**lan Cienki, **G**erard Steen, **G**raham Low, **L**ynne Cameron, **E**lena Semino, **J**oseph Grady, **A**lice Deignant and **Z**oltan Kövecses. The MIP comprises the following stages:

1. reading the whole text in order to establish its general meaning;
2. determining the lexical units that have been used in the text;
3. determining the meanings of the lexical units.

The procedure requires that every lexeme's meaning be determined in the given context, i.e. how the word "applies to an entity, relation, or attribute in the situation evoked by the text" (Pragglejaz 2007, p. 3). The next step involves determining whether each of the words has a different, more basic sense, activated in other contexts (e.g. *głowa kapusty* 'a head of cabbage' and *głowa* 'head' denoting a body part). According to the authors, more basic meanings of a lexeme include, cf.

1. concrete, physical meanings (what the word evokes is possible for the speaker to imagine; it can be seen, heard, felt, smelled, or tasted),
2. meanings evoking bodily action,
3. more precise, clear meanings,
4. historically older meanings

This does not imply, however, that a more basic meaning is also the most frequent one.

The MIPVU procedure (Steen *et al.* 2010) has been enriched by adding the following principles: word class boundaries may not be crossed by lexical units represented by the same shape (i.e. a meaning of a verb cannot be compared to a meaning of a noun), and word etymology is hardly ever considered. Additionally, metaphors have been divided into subtypes: *direct metaphors*, *implicit metaphors*, *personifications*, and there is a separate class grouping problematic expressions (*borderline cases of metaphor*). Metaphor signals appearing in the text are taken into account.

However, there are also doubts concerning this procedure, since too much attention is paid to data excerpted from dictionaries. It also takes great pains to establish "the more basic" meanings. In the SYNAMET project, the procedure of metaphor identification is slightly easier because the research field is restricted to synesthetic metaphors only.

5 Preliminary Guidelines for Synesthetic Metaphor Annotation in the Corpus

An annotation manual will be prepared, as well as a set of basic frames for the five perceptual domains — VISION, HEARING, TOUCH, TASTE, SMELL — together with a lexicon containing vocabulary that evokes each of the frames. Separate frames will be elaborated for the so-called *complex synesthesia*. In order to construct the specific frames, we will utilize the studies devoted to various aspects of perception: seeing (Dobaczewski 2002, Zawisławska 2004, Zielińska 2011, Dyszak 1999, 2010, Tokarski 2004), hearing (Żurowski 2012, Kładoczny 2012), smell (Bugajski 2004, Badyda 2013), touch (Bronikowska 2007) and taste (Mitrenga 2009, 2010, 2011). Both the frames and the lexicon will be elaborated in the most comprehensive and meticulous manner possible. However, since it is conceivable that an important element of a frame could be overlooked, the annotators will be granted the right to add the necessary elements during the corpus annotation.

The annotation of texts in the corpus will be performed on a hierarchical basis — firstly, one person (the annotator) analyzes the text, then a second person (the superannotator) checks the same text and introduces any necessary corrections. The fundamental changes will be those that indicate a difference of opinion between the annotators. It will enable us to distinguish between the more typical, central metaphors, and the less typical, peripheral ones.

The annotators will single out those excerpts where forms representing the perceptual lexicon enter into syntactic relations. They will also check whether the heads or the dependents of these phrases evoke the same perceptual frame (a literal expression, e.g. *śmierdziało lawendą* 'it stank of lavender'), a different perceptual frame (a metaphorical expression, e.g. *pachniał donośnie* 'it was smelling loudly'), or else a completely different conceptual frame (a metaphorical expression, e.g. *miał czarne myśli* 'he's been having dark thoughts').

Having identified the metaphorical expressions, annotators will then proceed to the *metaphor clusters* (MC) recognition. The term *metaphor cluster* refers to a phrase, a larger piece of text, or even a whole text centered around the same referent (e.g. olfactory sensation). Cf.

> *Encens Mythique d'Orient dość szybko zmierza do **ambrowego akordu bazy. A ta jest cudownie ciepła, głęboka, słodkawa i jednocześnie piżmowa, przy tym bardzo zmysłowa. Pachnie zaskakująco donośnie i bardzo długo — otaczając noszącego prawdziwie luksusową aurą.***
>
> ['Encens Mysthique d'Orient makes its way towards the amber accord of the base quite quickly. And the latter is wonderfully warm, deep, sweetish and musk at the same time. What's more, it's really sensual. It smells surprisingly loudly and very long, surrounding its user with an air of true luxury.'].

According to the above-mentioned assumptions, a substantial fragment of the text: *ambrowego akordu **bazy**. A ta [**baza**] jest cudownie ciepła, głęboka, słodkawa i jednocześnie piżmowa, przy tym bardzo zmysłowa. Pachnie zaskakująco donośnie i bardzo długo — otaczając noszącego prawdziwie luksusową aurą* constitutes a single metaphor cluster, since its referent is smell, evoked by the lexical unit *baza (zapachu)* '(fragrance) base.' As can be seen, a metaphor cluster does not necessarily coincide with the borders of the utterance.

Within a single MC we may find just one *metaphorical unit* (MU), several units, or even over a dozen MUs. Metaphorical units are defined as word forms or phrases which are used metaphorically in a given context, i.e. they combine lexemes primarily belonging to different perceptual frames, or to other types of frames. For instance, in the MC presented above, the following MUs may be distinguished:

- *baza* (*zapachu*) '(fragrance) base' [VISION and SMELL]
- *ambrowy akord* 'amber accord' [HEARING and SMELL]
- [*baza*] *jest cudownie ciepła* '[the base] is wonderfully warm' [TOUCH and SMELL]
- [*baza*] *jest cudownie głęboka* '[the base] is wonderfully deep' [SMELL and COMPLEX SYNES-THESIA]
- [*baza*] *jest cudownie słodkawa* '[the base] is wonderfully sweetish' [SMELL and TASTE]
- [*baza*] *pachnie zaskakująco donośnie* '[the base] smells surprisingly loudly' [SMELL and HEAR-ING]

First of all, the annotator will have to complete the elliptic MUs with the missing components, indicate the topic and the modifier, and describe their grammatical properties. Afterwards, the perceptual frames (or other types of frames) connected with the specific metaphorical expression will be chosen: a topic frame and a modifier frame.

For instance, the expression *baza* (*zapachu*) '(fragrance) base' is elliptical in that its topic is missing. Its primary meaning is 'a foundation, a bottom part.' In the text, the word is used as a perfumery term — *baza zapachu* — i.e. the last fragrance component to become perceptible. Hence, in that metaphor, *zapach* 'smell' (topic) is conceived of in terms of the visual perception (*baza* 'base' — modifier). In contrast, the expression *ambrowy akord* 'amber accord' links the SMELL frame with the HEARING frame. The form *ambrowy* serves as the topic (which activates the SMELL frame, and within the frame — the element OLFACTORY SENSATION), and the word *akord* is the modifier that evokes the HEARING frame.

The example examined above demonstrates the necessity to account for the "pre-metaphors" — lexicalized terms whose metaphorical origin is perceptible, and which conceptualize fragrance as an entity composed of many elements, e.g. *nuta głowy* 'head note,' *nuta serca* 'heart note,' *nuta bazy* 'base note,' *baza zapachu* 'fragrance base,' etc. The pre-metaphors often become the point of departure for more creative textual metaphors, e.g. **nuta** [zapachowa] '[fragrance] note' → *ambrowy* **akord** *bazy* ['amber accord of the base'] → [baza] *pachnie zaskakująco* **donośnie** 'the base smells surprisingly loudly'.

The annotators will classify the MU under the appropriate semantic category of metaphors. The categories' set shall be formed during the preliminary analysis of the samples. The categories that emerge at this stage include e.g. strong/weak synesthetic metaphor, simple/complex synesthesia. A separate QUASI_METAPHOR category is planned for the expressions whose metaphoricity is doubtful to the annotator.

The outcome of the MU analysis will comprise:

- grammatical categorization of the MU,
- a general metaphorical scheme of the MU, e.g. HEARING → SMELL,
- lexical items activating the MU's frame together with their grammatical description, i.e., for instance a combination of a lexeme primarily activating the HEARING frame with lexemes form the SMELL frame. It will establish the productivity of each of the lexemes.
- a detailed metaphor scheme of the MU, e.g. OLFACTORY SENSATION IS A SET OF SOUNDS,
- the MU semantic categorization, e.g. strong synesthetic metaphor, simple synesthesia.
- the tags for the metaphor signals.

The finished corpus will allow queries concerning specific lexical items and their metaphoric productivity, e.g.

słodki 'sweet'

→ *dźwięk* 'sound' → AUDITORY SENSATION → HEARING,
→ *zapach* 'smell' → OLFACTORY SENSATION → SMELL,
→ *kolor* 'color' → VISUAL SENSATION → SEEING.

Additionally, it will be possible to search for the perceptual frames, and to check how frequent the lexemes that activate the become either topics or modifiers, e.g.

→ modifier: *jasny dźwięk* 'light sound' (SEEING → HEARING),

→ topic: *ciepły/ciemny zapach* 'warm/dark smell' (SMELL ← TOUCH; ← SEEING).

The frame elements browser will enable users to check which elements of a specific frame (e.g. the HEARING frame) are associated with elements of a different frame (e.g. SEEING, SMELL, etc.), and to see the specific MU in which the connection manifests itself. For instance, the INTENSIVITY OF A VISUAL SENSATION describes the INTENSIVITY OF AN OLFACTORY SENSATION, cf. *niewyraźny zapach* 'blurry smell.'

Statistical analysis of the results will be performed in order to formulate general conclusions about the systematicity of synesthetic metaphors in Polish. The completed corpus will be accessible online to all registered users (free registration), and the results of the analyses are to be presented in a monograph written in English. It will also be possible to employ the corpus in other research projects.

References

Arutjunowa, N.D. (1981). *Metafora językowa (II) (Składnia i leksyka). Teksty.* 1 (55), 138–153.

Badyda, E. (2013). *"Upadły anioł zmysłów"?: metaforyka zapachu i percepcji węchowej we współczesnej polszczyźnie.* Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Biłas-Pleszak, E. (2007). "Zobaczyć dźwięk" — metafory synestezyjne jako przykład „korespondencji zmysłów". *Język Artystyczny.* 13, 157–166.

Black, M. (1971). Metafora. *Pamiętnik Literacki.* LXII. Z. 3, 217–234.

Black, M. (1993). More about metaphor. *Metaphor and Thought.* Ed. by A. Ortony. Cambridge: Cambridge University Press, 19–41.

Bogusławski, A. (1971). O metaforze. *Pamiętnik Literacki.* LXII. Z. 4, 113–126.

Bronikowska, R. (2007). Przymiotniki oznaczające cechy przedmiotów odbierane zmysłem dotyku we współczesnej polszczyźnie. *Studia nad współczesną polszczyzną. Gramatyka, semantyka, pragmatyka.* Red. A. Dobaczewski. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, 241–250.

Bugajski, M. (2004). *Jak pachnie rezeda. Lingwistyczne studium zapachów.* Wrocław: Atut.

Buttler, D. (1978). *Rozwój semantyczny wyrazów polskich.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Cameron, L. (2003). *Metaphor in Educational Discourse.* London & New York: Continuum.

Cameron L. & Low G. (Eds). (1999). *Researching and Applying Metaphor.* Cambridge: Cambridge University Press.

Cameron, L. & Maslen, R. (2010). Identyfying metaphors in discourse data. *Metaphor Analysis: Research Practice in Applied Linguistics, Social Sciences and the Humanities.* Eds. L. Cameron, R. Maslen. London: Equinox.

Classens, C. (1993). *Worlds of Sense.* London and New York: Routledge.

Cockiewicz, W. (2011). *Metaforyka Leśmiana. Analiza lingwistyczna.* Kraków: Księgarnia Akademicka.

Coulson, S. (2001). *Semantic Leaps. Frame-Shifting and Conceptual Blending in Meaning.* Cambridge: Cambridge University Press.

Deignan, A. (2005). *Metaphor and Corpus Linguistics.* Amsterdam/Philadelphia: John Benjamins.

Deignan, A. & Semino E. (2010). Corpus techniques for metaphor analysis. *Metaphor Analysis: Research Practice in Applied Linguistics, Social Sciences and the Humanities.* Eds. L. Cameron, R. Maslen. London: Equinox.

Dobaczewski, A. (2002). *Zjawiska percepcji wzrokowej: studium semantyczne.* Warszawa: UW. KLF.

Dobrzyńska, T. (1984). *Metafora.* Wrocław – Warszawa – Kraków – Gdańsk – Łódź: Zakład Narodowy im. Ossolińskich.

Dobrzyńska, T. (1994). *Mówiąc przenośnie... Studia o metaforze.* Warszawa: IBL.

Dyszak, A. (1999). *Językowe wyrażenia zjawisk emisji światła.* Bydgoszcz: Wydaw. Uczelniane WSP.

Dyszak, A. (2010). *Językowe wyrażenia zjawisk jasności i ciemności.* Bydgoszcz: Wydawnictwo Uniwersytetu Kazimierza Wielkiego.

Fass, D. (1991). met*: A method for discriminating. *Computational Linguistics.* 17(1), 49–90.

Fauconnier, G. & Turner, M. (2002). *The Way We think. Conceptual Blendings and the Mind's Hidden Complexities.* New York: Basic Books.

Magdalena Anna Zawisławska   - 117 -

SYNAMET — A Microcorpus of Synesthetic Metaphors. Preliminary Premises of the Description...

Fillmore, Ch. (1982). Frame Semantics. *Lingusitics in the Morning Calm.* The Linguistics Society of Korea (eds.). Seoul: Hanshin Publishing Co., 111–137.

Glucksberg, S. & Keysar, B. (1993). How metaphor works. *Metaphor and Thought.* Ed. by A. Ortony. Cambridge: Cambridge University Press, 401–424.

Goatly, A. (1997). *The Language of Metaphors.* London and New York: Routledge.

Goatly, A. (2006). Ideology and metaphor. *English Today.* 22(3), 25–39.

Grzesiak, R. (1983). *Semantyka i składnia czasowników percepcji zmysłowej.* Wrocław: Zakład Narodowy im. Ossolińskich.

Judycka, I. (1963). Synestezja w rozwoju znaczeniowym wyrazów. *Prace Filologiczne.* XVIII, 59–78.

Kładoczny, P. (2012). *Semantyka nazw dźwięków w języku polskim.* T.1–2. Łask: Oficyna Wydawnicza Leksem.

Knudsen, S. (2003). Scientific metaphors going public. *Journal of Pragmatics.* 35, 1247–1263.

Kövecses, Z. (2011). *Język, umysł, kultura. Praktyczne wprowadzenie.* Kraków: Universitas.

Krishnakumaran, S. & Zhu, X. (2007). Hunting elusive metaphors using lexical resources. *Proceedings of the Workshop on Computational Approaches to Figurative Language.* Rochester, NY: Anna Feldman, Xiaofei Lu, 13–20.

Lakoff, G. & Johnson, M. (1988). *Metafory w naszym życiu.* Warszawa: PIW.

Lundmark, C. (2005). *Metaphor and Creativity in British Magazine Advertising.* Luleå: Luleå University of Technology.

Minsky, M. (1975). A Framework for Representing Knowledge. *The Psychology of Computer Vision.* (Ed.) P. Winston. New York: McGraw-Hill, 211–277.

Mitrenga, B. (2009). Nazwy zmysłu smaku w języku polskim. *LingVaria.* R. 4. Nr 2, 227–236.

Mitrenga, B. (2010). Intensyfikowanie doznań smakowych w polskiej leksyce i frazeologii. *Ilość – wielkość – wartość.* Red. E. Umińska-Tytoń. Łódź: Archidiecezjlane Wydawnictwo Łódzkie, 317–328.

Mitrenga, B. (2011). Czasowniki percepcji smakowej w polszczyźnie historycznej i współczesnej. *Idea przemiany. Zagadnienia literatury, kultury, języka i edukacji.* T. 3. Częstochowa: Wydawnictwo WSL, 235–245.

Müller, C. (2008). *Metaphors Dead and Alive, Sleeping and Waking: A Dynamic View.* Chicago: University Of Chicago Press.

Najdecka, A. (2013). Innowacje semantyczne w nazwach kosmetyków. *Poradnik Językowy.* 4, 75–86.

Pragglejaz Group (2007). MIP: a method for identifying metaphorically used words in discourse. *Metaphor and Symbol.* 22(1), 1–39.

Prochowicz, A. (2013). Jak mówimy o śpiewaniu? Metafory synestezyjne jako element językowego obrazu śpiewu ludzkiego w polszczyźnie. *Linguarum Silva.* 2, 55–70.

Richards, I. (1936). *The Philosophy of Rethoric.* New York and London: Oxford University Press.

Ritchie, D. (2006). *Context and Connection in Metaphor.* New York: Palgrave-Macmillan.

Ritchie, D. (2011). "Justice is blind": A model for analyzing metaphor transformations and narratives in actual discourse. *Metaphor and the Social World.* 1, 70–89.

Rogowska, A. (2007). *Synestezja.* Opole: Oficyna Wydawnicza.

Rosińska, A. (2005). Rola synestezji w obrazowaniu zapachu. *Studia Filologiczne Akademii Świętokrzyskiej.* 18, 65–77.

Schank, R. & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Searle, J. R. (1993). *Metaphor. Metaphor and Thought.* Ed. by A. Ortony. Cambridge: Cambridge Uniwersity Press, 83–111.

Sedivy, S. (1997). Metaphors Picture, Pulsars, Platypuses. *Metaphor and Symbol.* 12 (2), 95–112.

Semino, E. (2008). *Metaphor in Discourse.* Cambdridge: Cambridge University Press.

Shutova, E., Lin, S., & Korhonen, A. (2010). Metaphor Identification Using Verb and Noun Clustering. *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010).* Beijing: Chu-Ren Huang and Dan Jurafsky.

Skorczynska, H. & Deignan, A. (2006). Readership and purpose in the choice of economics metaphors. *Metaphor and Symbol.* 21(2), 87–104.

Soskice, J. (1985). *Metaphor and Religious Language.* Oxford: Oxford University Press.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., Pasma, & T. (2010). *A Method for Linguistic Metaphor Identification. From MIP to MIPVU.* Amsterdam: John Benjamins.

Stefanowitsch, A. & Gries S. Th. (Eds). (2006). *Corpus-Based Approaches to Metaphor and Metonymy.* Berlin: Mouton de Gruyter.

Tannen, D. (1993). What's in the Frame? Surface Evidence for Underlying Expectations. *Framing in Discourse.* Ed by D. Tannen. New York: Oxford University Press, 14–43.

Termińska, K. (1992). Metafora synestezyjna. *Poradnik Językowy.* 3, 201–207.

Tokarski, R. (2004). *Semantyka barw we współczesnej polszczyźnie.* Lublin: Wydaw. Uniwersytetu Marii Curie-Skłodowskiej.

Ullman, S. (1962). *Semantics: an introduction to the science of meaning.* Oxford: Blackwell.

Werning, M., Fleischhauer, J., & Beseoglu H. (2006). The cognitive accessibility of synesthetic metaphors. *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society.* (Eds.) R. Sun & N. Miyake. London: Lawrence Erlbaum Associates, 2365–2370.

Wierzbicka, A. (1971). Porównanie – gradacja – metafora. *Pamiętnik Literacki.* LXII. Z.4, 126–147.

Witucka, M. (1998). Jak opisuje się zapachy w reklamie perfum. *Poradnik Językowy.* 3, 1–8.

Wróblewski, P. (1998). *Struktura, typologia i frekwencja polskich metafor.* Białystok: Wydawnictwo Uniwersytetu w Białymstoku.

Zawisławska, M. (2004). *Czasowniki oznaczające percepcję wzrokową we współczesnej polszczyźnie: ujęcie kognitywne.* Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Zielińska, K. (2011). *Obiekt w (semantycznym) polu widzenia: analiza kontrastywna czasowników percepcji wzrokowej w języku polskim i niemieckim.* Warszawa: Uniwersytet Warszawski. Instytut Germanistyki. Wydział Neofilologii.

Żurowski, S. (2012). *Wyrażenia percepcji słuchowej w języku polskim: analiza semantyczna.* Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Publisher: Institute of Slavic Studies, Polish Academy of Sciences, University of Silesia & The Slavic Foundation