

Jan Wieczorek  
Uniwersytet Wrocławski

## **The Semantics of Words Comprising Phraseological Structures – the Observations of Polish Corpus of Wrocław University of Technology Editor**

As a distinct discipline of lexicology, phraseology recognizes the phrasemes as “socially fixed word combinations that are irregular in some respect.”<sup>1</sup> In consequence of the said irregularity the phraseme has a certain autonomy in both the structural and the semantic field. In the former case, the phraseme should be treated as a fixed entity with collocational restriction (or with a specific collocational pattern). The entity’s structure is either regular or susceptible to small changes that adapt the phraseme to be used in a text. In the latter case, the phraseme also has a distinct meaning, which is not a sum of partial meanings, but something additional. Usually, the object of semantic research is a phraseme as an entity, not the words that amount to it. However, sometimes the editor of linguistic corpora (here: the Polish Corpus of Wrocław University of Technology) encounters such cases that require careful observation and analysis of the meanings of distinct words that make up various fixed phrases.

The material used in this work was gathered during WSD (Word Sense Disambiguation) annotation. It is a process of prescribing the meaning to highlighted homonyms which appear in the corpus’ texts. The initial set of definitions of lexemes chosen for annotation is based on the descriptions of their meanings in plWordNet (Polish WordNet, a specific digital dictionary). When it comes to ordinary words, which are used only in text collocations, the process (that is the selection of a matching definition from a list for each word appearance in the texts) runs smoothly. It has been observed, though, that some words are difficult to annotate: parts of metaphors, named phrases, terms originating

---

<sup>1</sup> Pajdzińska A., Lewicki A.M., “Frazeologia” [in:] *Współczesny język polski*, ed. Bartmiński J., Lublin 2001, p. 315.

from sociolects (other nonstandard usage), and phrasemes. New definitions had to be added to the list of meanings. Below there is an example of a complete definition set, which is used by the expert linguist performing the WSD annotation. The set was prepared by one of the linguists of the WrocUT Language Technology Group. It is a complete set of dictionary entries (or rather entries recognised as basic by the automatic morphological analyser) for the Polish word *myszka* [a little mouse].

myszka-1

a small or pleasant mouse (“He’s also keen on animals and keeps a *myszka* at home.”)

myszka-2

a birthmark, small, brownish, hairy spot on a person’s skin (“If it’s really her, she should have a crescent shaped mark, the so-called *myszka*.”)

myszka-3

a colloquial and humorous name of vulva (“Also her *myszka* was not fully developed – only a small pink slit with barely there clitoris.”)

myszka-4

an electrical device used for handling a computer, which allows the cursor to move on the screen without using the keyboard (“Today Logitech announced the launch of a new *myszka*.”)

myszka-5

sometimes an idiolectic name of a girl or woman (“Her mother calls her *Myszka*, but her real name is Marysia.”)

myszka-6

0.4 – *szara myszka* [a grey mouse] a woman that is modest, unassuming, quiet.

myszka-7

0.4 – *trącić myszką* [to smell like a mouse] something that is outdated, stale, old (initially used to describe the savour of wine)

myszka-90

FOREIGN WORD

myszka-91

TAGGER ERROR

myszka-92

NAMED PHRASE

myszka-93  
OTHER MEANING

myszka-99  
FAULTY TEXT

To completely describe the meanings in the corpus, it was necessary to define the meanings of all words that make up phrasemes. Sometimes new meanings had to be ascribed, when the word sense used in a non-standard way could not be matched to previously established definitions (I would like to point out that among the meanings of *myszka* a new sense of the words is created in the phraseme *trącić myszką*, but in the phrase *zabawa w kotka i myszkę* [to play like a cat and a mouse] the sense brings us closer to reality – to an actual mouse – and thus can be annotated using the first of the available definitions). Those definitions have a 0.4 tag. The definition of such a word can be vague and partial, but should be clear to the expert linguist and as short as possible. Those words are defined in an experimental and abstract way, since the more or less precisely established meaning does not appear separately or probably in any phrase other than the one being analysed. Therefore, such meaning is purely potential – even in a given phraseme it is secondary to the general meaning of the phraseme, which is definitely more important for communicating the message. The lexicographical tradition supports this statement.

This work aims at presenting the phrasemes analysed for the corpus and their definitions proposed by the expert linguist for annotation purposes. Also, I would like to make a preliminary analysis of the material to examine the possibility to define entities with dependent meaning, that is words making up phrasemes, as well as to recreate the expert linguist's mindset and justify some of the choices made.

## 1. Review and classification of the material

Below I present the analysed phraseological material as keywords (lexemes that were of interest to the group of expert linguists during the corpus' annotation) in alphabetical order that appear in respective phrasemes. Each entry contains a keyword and an approximate explication of the words use in a phraseme or a group of phrasemes. Unfortunately, the gathered material is sparse and not diversified enough to make it reliable, but sufficient for a semantic analysis, which attempts to justify itself. For every 89 lexemes (nouns and verbs) chosen for annotation, 11 have in their set of definitions at least one that tries to convey a specific meaning of the word that is a part of a phraseme.

dać-20 [to give]

0.4 – *dać radę* – to manage, cope, meet the challenge.

dać-21

0.4 – *dajmy na to* – to assume, give an example.

język-11 [tongue]

0.4 – *być na językach* [to be on people's tongues], *złośliwe języki* [spiteful tongues] – phrasemes about gossiping.

mówić-11 [to speak]

0.4 – *nawiasem mówiąc* [speaking in parenthesis], *mówiąc jakoś* [honestly, actually speaking] – to mention something unofficially, incidentally, to agree with others on the way things are.

myszka-6 [a little mouse]

0.4 – *szara myszka* [a grey mouse] – a woman that is modest, unassuming, quiet.

myszka-7

0.4 – *trącić myszką* [to smell like a mouse] – something that is outdated, stale, old (initially used to describe the savour of wine).

oko-18 [an eye]

0.4 – *na oko* [by the eye], *rzut oka* [to throw an eye on something], *mieć oko* [to have an eye] – a group of phrasemes which emphasise the quality of having a sharp eye, being perceptive, correctly estimating the quantity.

oko-19

0.4 – *na oczach* [in front of the eyes], *w (żywe) oczy* [at lively eyes] – a group of phrasemes referring to the eye as an organ of sight, emphasising their direct involvement in the event, communication or observation.

oko-20

0.4 – *bić/kłuć/rzucać się w oczy* [to hit/stab/throw yourself at eyes] or *dawać po oczach* [to give in the eyes] – to be striking, glowing, seeking attention.

oko-21

0.4 – *w oczach* [in the eyes] – fast, clearly.

oko-22

0.4 – *otworzyć komuś oczy na coś* [to open somebody's eyes onto something], *w czyichś oczach* [in somebody's eyes], *patrzeć jakimś okiem* [to look with a certain eye] – a group of phrasemes which emphasise the subjectivity of sensual experience.

oko-23

0.4 – *bez mrugnnięcia/zmrużenia oka* [without a blink] – in cold blood.

oko-24

0.4 – *z przymrużeniem oka* [with a wink] – not serious.

oko-25

0.4 – *w mgnieniu oka* [in a blink of an eye] – quickly.

piłka-5 [a ball]

0.4 – *krótka piłka* [a short ball] – a quick question or errand.

piłka-6

0.4 – *piłka po czyjejs stronie* [a ball on somebody's side], *zagrać piłkę do kogoś* [play a ball to somebody] – waiting for somebody's response in a discussion or for a reaction to an act.

położyć-10 [to lay]

0.4 – *położyć krechę* [to lay a line], *położyć krzyżyk* [to lay a cross], *położyć kres* [to lay an end], *położyć lachę* [to lay a stick], etc. – meanings referring to setting a limit to something, or to dissociate one thing from another.

położyć-11

0.4 – *położyć akcent* [to lay an accent], *położyć nacisk* [to lay a pressure] – meanings referring to emphasising something.

posiadać-30 [to possess]

SIE [reflexive] 0.4 – *nie posiadać się* [not to have oneself] – meaning referring to self-control [Janek *nie posiadał się* from happiness after hearing about your arrival.]

prosić-11 [to ask, implore]

0.4 – *proszę państwa* [I ask You (plural)], *proszę cioci* [I ask you, Aunt] – a polite form of address, meaning is difficult to grasp.

ręka-8 [a hand]

0.4 – *robić coś na własną rękę* [to do something on one's hand], *wolna ręka* [a free hand] – meaning referring to an independence, a freedom of choice.

ręka-9

0.4 – *z drugiej ręki* [from the second hand] – with other's mediation, also an antonym to *z pierwszej ręki* [from the first hand].

ręka-10

0.4 – *iść komuś na rękę* [to go onto one's hand], *być na rękę* [to be on a hand] – consistent with one's plans.

ręka-11

0.4 – *od ręki* [from a hand], *pod ręką* [under a hand], *w zasięgu ręki* [in a hand's reach] – immediately available.

ręka-12

0.4 – *być w czyichś rękach* [to be in one's hands], *położyć na czymś rękę* [to lay a hand on something], *wpaść w ręce kogoś* [to fall into somebody's hands], *dostać coś z rąk kogoś* [to get something from one's hands] – meanings referring to power.

ręka-13

0.4 – *dostawać/zarabiać na rękę* [to get/earn on a hand] – a salary after taxes, net income.

ręka-14

0.4 – *złożyć coś na ręce kogoś* [to lay something on someone's hands], *umywać ręce* [to wash one's hands], *brać w swoje ręce* [to take something in one's hands] – meanings referring to one's responsibility for something or for doing something.

ręka-15

0.5 – *pomocna ręka* [a helpful hand], *wyciągnąć rękę* [to reach out the hand] – a help, a goodwill, an agreement.

związać-12 [to bind]

0.4 – *związać koniec z końcem* [to bind together two ends] – to have enough money to make a living.

związać-13

0.4 – *związać komuś ręce* [to bind someone's hands] – to prevent someone's action.

związać-14

0.4 – *związać ręce stulą* [to bind hands with a stole] – to marry a couple.

From a functional standpoint, the presented phrasemes can be classified as follows:

- a) phrases: *dać radę*, *dajmy na to*, *nawiasem mówiąc*, *krótka piłka*;
- b) expressions: *mieć mocną głowę*, *być na językach*, *trącić myszką*, *mieć oko*, *bić/kłuć/rzucać się w oczy*, *dawać po oczach*, *otworzyć komuś oczy na coś*,

*patrzeć jakimś okiem, zagrać piłkę do kogoś, położyć krechę, położyć krzyżyk, położyć kres, położyć lachę, położyć akcent, położyć nacisk, prowadzić za rękę, prowadzić na pasku, prowadzić jak stado baranów, iść komuś na rękę, być na rękę, być w czyichś rękach, położyć na czymś rękę, wpaść w ręce kogoś, dostawać/zarabiać na rękę, złożyć coś na ręce kogoś, umywać ręce, brać w swoje ręce, wyciągnąć rękę, związać koniec z końcem, związać komuś ręce, związać ręce stulą;*

c) nominal phrases: *złośliwe (kłamlive) języki, szara myszka, piłka po czyjejś stronie, rzut oka, pomocna ręka;*

d) modifier phrases: *na oko, na oczach, w (żywe) oczy, oko w oko, bez mrugnienia/zmrużenia oka, w mgnieniu oka, od ręki, pod ręką, w zasięgu ręki, mówiąc jakoś (szczerze, konkretnie).*

As this classification implies, the group that most often required establishing new meanings (in the aforementioned annotated sample) were the expressions with nominal phrases and modifier phrases being much rarer. In the presented set there were no adverbial phrases, which could be due to the fact that the annotated words were only nouns or verbs; there were no conjunctions, particles or operators that are parts of adverbial phrases.

## 2. Semantic analysis of words appearing in phrasemes

As the phraseme is defined as a word combination that is irregular in some respect, it is worth to consider how big the irregularities are in the aforementioned phrasemes – are they idioms, or rather compositional phrasemes (a combination of words with meaning encompassed by its semantic pivot). Also, we must decide, whether the irregularities that span the whole phraseme modify the meaning of its member lexemes in a substantial way. If so, then to what extent and is it possible to distinguish and describe more general tendencies?

Even in this modest sample there is an apparent tendency for the names of body parts to participate in phrasemes. In the process, the body part is usually associated with its everyday function, or impact on the organism (e.g., *serce* [a heart] – a word not included in the list – the central element of an organism, and also, metaphorically, of a machine, e.g. *sercem statku jest maszynownia* [the heart of a ship is its engine room]). Such tendency stems from the anthropocentrism of the language, a natural process of describing the world in terms which are already known and used. There are no meanings better known than the ones that describe the human body. Similar metaphors are often textually realised in phrasemes. Since the metaphors' meanings do not diverge much from the

prototypical ones, or sometimes even the primary ones, the phrasemes should be considered the compositional phrasemes rather than idioms.

The aforementioned phenomenon of metaphorising body parts – which under the cultural convention refer to their associated qualities and functions – will be observed in further phrasemes from the corpus. An excellent example is the entity *być/znaleźć się na językach*, where the lexeme *język* (colloquially) refers to the body part responsible for talking. Obviously, the meaning of the phraseme is not entirely consistent with the meaning of each of its member lexemes. The proper explication is: *być przedmiotem plotek* [to be gossiped about], and not *znaleźć się (fizycznie) na czyimś języku* [to (physically) appear on someone's tongue]. It should be noticed, though, that in this case the semantic irregularity does not stem from nonprototypical meaning of *język* but rather from the convention referring to the composite phraseological entity. The researched lexeme is contained into linguistic prototype, which allows for this quite regular metaphor. Another example of metaphorising *język* is the expression *złośliwe/klamliwe języki* where the user of an organ of speech is associated with the organ<sup>2</sup>. The origin of both metaphors is *język* meaning a body part, and the target term is the product of the tongue usage (here: a gossip) and the tongue user (here: a gossiper). The difference between the presented phrasemes is, however, the extent of idiomatisation: *być/znaleźć się na językach* seems to be more conventionalised than *złośliwe/klamliwe języki*, a phraseme that is semantically clear.

A rich source of phrasemised metaphors are the lexical entities *oko* and *ręka*. The conceptualisation of *oko* is textually realised in different yet related meanings. In all meanings the metaphorical origin is *oko* as an organ of sight. In the first group of meanings (*na oko, rzut oka, mieć oko*) the lexeme refers to the quality of having a sharp eye, being perceptive, correctly estimating the quantity. In contrast to the former examples (*głowa, język*), it is much more difficult to assign the metaphor to the meanings closer to the prototypical meaning of *oko*. The number of metaphor levels suggests that the meanings are peripheral, nonprototypical or singular. The metaphor is obscured, although it can be intuitively understood.

The second group of phrasemes containing the lexeme *oko* is a semantically consistent group of: *na oczach, w żywe oczy, oko w oko*. The meaning of these phrasemes refers to the direct involvement in the observation, communication or participation in some events. Here *oko* also refers to the organ of sight, but this time the sensual function has no influence on the resultant metaphor. The

---

<sup>2</sup> The realisation of this metonymy is also an expression *język* referring to an informer, e.g., a captive interrogated for enemy actions. The meaning is lexicalized and has its own definition.



individual, subjective and direct character of perception is accented; to see something *własnymi oczami* [with one's own eyes] one must be in a certain place in person and within short range. What is important here is not the sight, but the context of its usage. The expression *w oczach* (e.g., Jerzy was growing *w oczach*, Krzysztof was withering *w oczach*) is similar. The lexeme *oko*, as in the previous example, refers to the eye meaning an organ of sight, though neither the sight function nor the context are important. The metaphor results from a temporal feature of perceiving something with the sense of sight. Usually, continuous observation of an object or a phenomenon is moderately short, as it requires full personal engagement of the observer, who is at the time unable to perform any other activity. While something is growing or withering *w oczach*, it means that it changes in a duration of a single observation or between one glance and another.

The fourth meaning of the lexeme *oko*, which appears in the phrasemes, refers to being striking, glowing, seeking attention: *bić/kłuć/rzucać się w oczy* and *dawać po oczach*. Here, the eye is again the organ of sight, more precisely the receptor exposed to sight sensations. The meaning of the phrasemes is rather clear and results from a juxtaposition of the sight receptor with the verbs meaning intense or aggressive influence (*bić, kłuć, rzucać się...*). It is possible that the connotations with the eye features and structure are important in the phrasemes. The eye is sensitive, delicate, and fragile, so combining it with verbs of intense action results in a vivid contrast.

In the phrasemes: *otworzyć komuś oczy na coś, w czyichś oczach, patrzeć jakimś okiem*, the sixth meaning of the lexeme *oko* can be found and it is still used a sight metaphor. This time the prominent sight feature is its subjectivity analogous to the subjectivity of one's judgements and opinions. In contrast with the three former groups, which could be analysed, the meaning of phrasemes in this group is hardly legible, and the language user unaccustomed to the convention can have problems with understanding the message. It means that these phrasemes are idiomatic, even though they contain the prototypical (and specific not only to the phrasemes) lexeme *oko*.

The aforementioned five meanings of the lexeme *oko* found in the five groups of phrasemes are semantically similar: they refer to an eye – an organ of sight. The meaning is quite frequent and is defined in plWordNet as:

oko-2

an ability to see (*kto ma dobre oczy, dojrzyć go może w dzień* [one who has good eyes, can see it in daylight], *chciałam zejść z oczu wiecznie wstawionej kobiecie* [I wanted to leave the eyes of the constantly intoxicated woman], *to ciekawe popatrzeć na własne dzieło cudzymi oczyma* [it is interesting to see one's doing with

someone else's eyes], *noc już była ciemna, że choć oko wykol* [it was pitch black, like with eyes put out], *bo mu się nawinęły pod oczy Dominikowa z Jagną* [because Dominikowa and Jagna came into sight right under his eyes"].

However, the modifications to the meaning of *oko* as part of a phraseme resulted in need of establishing new meanings, specific only to phrasemes (or to a small group of metaphors which are not phrasemes). In general, the modifications expose a feature of sight (represented in the text as the lexeme *oko*), that is often a vague connotation or nonverbal observation of everyday life.

The remaining two groups of phrasemes contain the word *oko*, but the meaning is *okolice oka* [eye's neighbourhood] or to be more precise – the eyelids. The phrasemes are: *bez mrugnięcia / zmrużenia oka* and *w mgnieniu oka*. The two illustrative phrasemes refer to the observed situations of blinking/winking. In the former phraseme, blinking is symptomatic to hesitation, abhorrence or aggravation; in the latter there is a link to the duration of blinking. As the dictionary definitions were insufficient to describe the function of the word *oko* in both phrasemes, the WSD editor has decided to add two new meanings: the first one is the lack of wink as a display of confident attitude, the second one is referring to blinking dynamics.

Here, I would like to go back to the aforementioned *myszka*. In the Polish Corpus of Wrocław University of Technology there are two phrasemes containing this lexeme: *szara myszka* and *trącić myszką*. The first phraseme is used to describe a modest, unassuming and quiet woman. The features are epitomised by a grey mouse, commonly found in rural areas. A small and hardly noticed, yet a very common animal, serves as a metaphor of a person with the same qualities. The second phraseme refers to the odour of a spoiled wine, described with the name *myszka*. Without this piece of information the origin and meaning of the phraseme *trącić myszką* cannot be deciphered: to have a smell of a wine (an old alcoholic beverage). The two examples have the same pivot and perfectly manifest different levels of semantic clarity. On the one hand, *szara myszka* is a very clear phraseme; one can easily decipher its meaning when one is familiar with Polish rural areas. On the other hand, the phraseme *trącić myszką* has kept its old meaning despite the lack of semantic motivation awareness (so it prevails on the basis of cultural convention). In both cases, the definition list of *myszka* from plWordNet had to be expanded, as the existing descriptions were inadequate to the encountered expressions.

All the previously discussed examples of phrasemes involve a noun as the annotated semantic pivot. To conclude this analysis, I would like to put forward a few cases, in which the keywords (from the WSD annotation perspective) are the verbs *dać* and *położyć*. *Dać* is present in the phrasemes *dać radę* (to manage,

cope, meet the challenge) and *dajmy na to* (to assume, give an example). In both cases, the meaning of the verb *dać* is far from the basic dictionary definitions present in the plWordNet. The recipient's lack of understanding of these idioms can render a message entirely incomprehensible. For example, the sequence *dać radę* can be interpreted literally, if one assumes an ordinary syntactic combination, such as giving somebody a piece of advice or instructions. Because of the lack of semantic motivation awareness of both phrasemes and a large difference of meaning, the corpus editor decided to establish two semantically vague meanings of the lexeme *dać*. The description of the terms is purely functional and associated with the phrasemes that contain the lexemes.

The verb *położyć* appears in the Polish Corpus of Wrocław University of Technology in two phraseme groups: *położyć krechę*, *położyć krzyżyk*, *położyć kres*, *położyć lachę* (meanings referring to setting a limit to something, or to dissociate from something) and *położyć akcent*, *położyć nacisk* (meanings referring to emphasising something). In the first case *położyć* can be associated with the basic meanings of the lexeme, for example:

*położyć-1*

to place something somewhere, on top of something, usually horizontally (he *położył* the book in an unobscured spot, so as not to forget it when leaving; he of course forgot it when he left)

and

*położyć-3*

to erect, build something (This spring the masons *położą* the foundation for the new theatre building. They *położyli* the tram rails at random.)

However, neither definition corresponds to the literal meaning of the phrasemes, which describe a metaphorical and abstract situation. The second group of phrases relates to writing a mark to emphasise something. This conceptualisation is illustratively treating an actual act of emphasising a given aspect of an errand, a meaning of a message, or a settlement condition. There is a meaning in plWordNet of writing a mark (*położyć 5 – napisać*, *położyć znak* [to write a mark]), but it suggests only the literal, physical interpretation.

### 3. Conclusion

The problem of analysing the meaning of words in phrasemes may seem artificial and unjustified in the general context of phraseology. Indeed, the

meaning of phrasemes—to a varying degree—depends on the meanings of the composite lexemes. Also, the linguists agree that phraseological entities should be treated as semantic entities, which can be subject to semantic analysis. Such perspective may seem to render this work ill-founded. However, it is worth to point out at the developing branch of language technology, where employees aim at creating information systems able to perform partially autonomous actions on the natural language. The programs with implemented artificial intelligence are already able to perform not only syntactic and morphologic analysis, but also first automatic summarizations of longer texts. The performance and the quality of these programs' outputs is largely dependent on the precision and functionality of the description of the language set on which the programs' artificial intelligence is trained to solve problems. The aforementioned WSD procedure aims to describe the means of disambiguating the meanings of polysemic entities. Sometimes, one of the meanings is a lexeme that participates in a phraseme, which implies the necessity of developing an explanation method of how unusual the encountered situation is to the artificial intelligence, and the methods of resolving this situation. This necessity has motivated the editors of the Polish Corpus of Wrocław University of Technology to undertake an original task – to establish the meanings of the entities which are parts of phrasemes. This work tries to retrace the reasoning and conclusions of these editors. At the same time they are conscious that the detailed description of this task requires more effort and exhaustive research techniques (e.g., atomic analysis of meanings). It is certainly not the last time the corpus editors ponder on the semantics of the word appearances in the text which are difficult to describe.

All the definitions fragments are extracted from the Inforex web-based system for annotating text corpora

*Tłumaczenie z języka polskiego na angielski: Julia Gryszczuk-Wicijowska*

Jan Wieczorek

**The Semantics of Words Comprising Phraseological Structures – the Observations of the Editor of the Corpus of Wrocław University of Technology**

The aim of the article is to present several problems encountered by the editor compiling the Corpus of Wrocław University of Technology. The problems are connected with the application of the method of disambiguation of meanings of words in text (WSD – Word Sense Disambiguation) during the work on material of phraseological

character. Any researcher using this method confronts the problem of the lack of semantic independence of words entering into the composition of fixed word combinations. The necessity to develop a unique procedure of conduct for this type of words, inclined the author of the article to experiment in the isolation of independent meanings of words co-forming phraseologisms and to make an attempt at their typologization. The isolation of such meanings has proved possible, although the accuracy of their description is very unstable: in some cases the motivation of semantic changes—which the meaning of a word undergoes under the influence of “being petrified” in a set phrase—is transparent, in other cases, however, we can speak of a hypothetical source of changes rather than a definite modification of meaning. It is worth mentioning here about the rigour imposed by the adopted WSD marking method: a set of all meanings which a given word form may assume in the analysed text must be entered into the database. That is why the problem of providing the definition of words forming phraseologisms is vital from the point of view of a linguist dealing with machine processing of language.

**KEYWORDS:** corpus linguistics, WSD, Word Sense Disambiguation, phraseology, semantics, lexicography, lexicology.

**mgr Jan Wiczorek** – doctoral student, Institute of Polish Philology at the University of Wrocław (UWr), promoter: Prof. Waldemar Żarski, Ph.D. (UWr). He is writing a dissertation devoted to the evolution of the conceptualization of notions TOWN and VILLAGE in the Polish language. Research assistant at CLARIN-PL Language Technology Centre at the Institute of Informatics of Wrocław University of Technology. Research interests: anthropological linguistics, semantic evolution, corpus linguistics.