MONIKA ZAŚKO-ZIELIŃSKA[1,A] & MACIEJ PIASECKI[2,B]

[1]Wrocław University, Poland
[2]G4.19 Research Group, Wrocław University of Technology, Poland
[A]monik@uni.wroc.pl ; [B]maciej.piasecki@pwr.edu.pl

LEXICAL MEANS IN COMMUNICATING EMOTION IN SUICIDE NOTES — ON THE BASIS OF THE POLISH CORPUS OF SUICIDE NOTES

Abstract

Polish Corpus of Suicide Notes (PCSN) is a relatively large set of authentic suicide notes that are linguistically annotated on several levels. In order to identify features characteristic for this genre we compared PCSN with the collected sub-corpus of counterfeited suicide notes. In this paper we focus on the lexical means of expressing emotions. Our goal was to analyse ways of expressing emotions in this specific genre. Our initial list of lexical markers was based on Markowski's list of the lexis common for different genres. The list was next expanded with the help of the plWordNet 2.0 — a lexico-semantic network. The expansion was based on the manually selected noun and verb hypernymy branches according to their correspondence to the elements of the initial list. For words from the extended list, a quantitative analysis was performed for both authentic and fake suicide notes. We have also analysed the use of the lexical markers of emotions, feelings and emotional states, as well as emotion operators, and ways of expressing personal evaluation, affection and hate.

Keywords: suicide note; corpus; lexical means for expressing emotions; sentiment analysis; Polish Corpus of Suicide Notes

1. The Polish Corpus of Suicide Notes

The *Polish Corpus of Suicide Notes* (henceforth PCSN) is a set suicide notes that were scanned and next manually transcribed. PCSN includes 614 genuine notes from the years 1998–2008 that were acquired from the public prosecutor's offices from the area of Poland. The collected texts have been written by 382 authors, where the youngest was 12 years old and the oldest 89 years. The collection includes:

- 456 suicide notes written by men (74.26% of all documents in the collection; written by the 290 distinct male writers),

- 158 suicide notes written by women (27.73% of all notes; written by 92 different women).

This ratio is close to the proportion observed in the Polish statistical data concerning committed suicides from the studied period, i.e. 77.78% male cases vs 22.21% female cases. PCSN includes also the *Subcorpus of Counterfeited Suicide Notes* (SCSN) which consist of 117 notes acquired during experiments in which experiment participants were asked to write a suicide note not in his/her own name, but on behalf of an imaginary person whose characteristic was presented to the participant as a part of the task description. In order to select the most representative data for the experiment (to generate tasks for participants), we had to collect data about people who had committed suicides in Poland in a way as complete as possible. We were especially interested in those who had left suicide notes. The data were extracted from both the available statistical data concerning the authors of suicide notes and people committing suicides in general, as well as the research results in this area (Zaśko-Zielińska, 2013). PCSN has been built as a result of the research project: "Suicide notes — linguistic methods of authorship attribution of text".

2. Goal

The main topic of this work are descriptions of emotions included in Polish suicide notes. Our aim is not recognition of the speaker's emotional state, but an analysis of the ways of expressing emotions in such a specific genre as suicide notes. We expected that the observed characteristic features would facilitate distinguishing between genuine and counterfeited notes, that is one of the goals of the forensic linguistics. Supporting of such a research, was one of the reasons to build PCSN.

In the work presented here, we focus on the lexical means of communicating emotions. We leave all other aspects apart as they require rather a separate work (e.g. syntactic-semantic means) or have already being described (e.g. graphical means including emotional punctuation) (Zaśko-Zielińska, 2013). We wanted to identify a list of words that can be used as lexical markers for the recognition of the emotional sentiment of the analysed letters. According to our initial assumption we have been looking for:

- names of emotions or affects and descriptions of the emotional states,
- emotion operators (i.e. dependant modifiers, see Sec. 5),
- emotion intensifiers,
- valuation words,
- *terms of endearment* and invectives.

3. Analysis Scheme

At the beginning the initial list of the lexical means for expressing emotions was constructed on the basis of *Markowski's list of the lexis common for different genres* (Markowski, 1992). The initial list included only one-word *lemmas*.[1] Next, we used

---

[1] A lemma is a basic morphological form and represents a set of word forms that have the same meaning and differ only in the value of the morphological categories, like number, gender, case etc.

the created list as a departure point to analyse ways of communicating emotions on the lexical level in the suicide notes.

In the next step, we tried to expand the list with lexical units[2] that are semantically associated in a way suggesting the similarity of their use in the context of communicating emotions. The list was extended on the basis of plWordNet 2.0[3] (Polish name: *Słowosieć*) (Maziarz, Piasecki, & Szpakowicz, 2012) that is briefly presented below. For the expansion we used the lexico-semantic relations of synonymy and hypernymy (Maziarz, Piasecki, & Szpakowicz, 2013). We assumed that lexical units which are linked by synonymy in plWordNet or belong together to the same, relatively small hypernymy subtree, share the emotional aspects of their lexical meanings with their synonyms and close hypernyms (i.e. inside a relatively small hypernymy subtree).

The acquired set of words,[4] in fact lemmas, have been compared with the lemma frequency extracted from PCSN and SCSN. Those frequencies, merged from both corpora, were used to build a frequency ranking of lemmas from the acquired set.

plWordNet is a lexical semantic network in which lexical units (representing lexical meanings) are described by lexico-semantic relations such that the given lexical unit participate in, e.g.

- lexical units are named according to the scheme: *lemma*, *sense identifier*,
- the lexical unit *samochód* 1 'a car' is described by:
- a set of lexical units linked to it by synonymy: {*samochód* 1 'a car', *pojazd samochodowy* 1, *auto* 1 'auto', *wóz* 1 '≈a car'} 'motor vehicle, automotive vehicle',
- its hyponyms: {autobus 1} 'bus, autobus, coach', {samochód ciężarowy 1, ciężarówka 1, ciężarówa 1} 'truck, motortruck', {samochód osobowy 1, osobówka 1, samochód 2, auto 2} 'car, auto, automobile, machine, motorcar', . . .
- its hypernyms: {dwuślad 1, pojazd dwuśladowy 1 dwuśladowiec 1} 'vehicle' is a hyponym of {pojazd 1}'≈vehicle' is a hyponym of {pojazd 2} '≈vehicle, craft' is a hyponym of { środek lokomocji 1, środek transportu 1} 'conveyance, transport'. . .
- its meronyms: { podwozie 1} 'chassis', { karoseria 2} 'car body', {deska rozdzielcza 1, tablica rozdzielcza 1} 'dashboard, fascia', {bagażnik 1} 'luggage compartment, automobile trunk, trunk, boot', {kierownica 1} 'steering wheel, wheel', {tłumik 2} '≈regulator (in a car)', . . .
- its holonyms: { pociąg drogowy 1} '≈truck, motortruck'
- and the lexical units linked to it by the inter-register synonymy: {rakieta 4} '≈car', {fura 4, bryka 3, autko 1} '≈car'

The plWordNet project is an attempt to apply the relational paradigm of lexical semantics on a very large scale. It is also aimed at covering a set of lemmas which is

---

[2]Following the model of plWordNet, a lexical unit is defined here as a pair: lemma plus lexical meaning identifier.

[3]http://plwordnet.pwr.edu.pl

[4]The whole set is too big to be presented in this paper. Instead, we have published this set and all other lists created for this research on the homepage of PCSN: www.pcsn.uni.wroc.pl.

larger than in any other contemporary dictionary of Polish. The most contemporary version (2.3 from the July 2015) includes:

- 170 834 lemmas, including 32 517 polysemous lemmas,
- 244 286 lexical units representing different lexical meanings,
- and more than 600 000 instances of the lexico-semantic relations (about 40 main types of relations and more than 90 subtypes).

Different plWordNet hypernymic subtrees can encompass hundreds, and in some cases even thousands, lexical units. The exact number depends on the particular thematic domain and how general is a hypernym which is selected as the top root node for the whole hypernymic subtree. Hypernymic subtrees include also sets of synonyms, as it can be observed in Figure 1, e.g. the lexical unit *uczucie* 5 'emotion' belongs to the set of synonyms {*emocja* 1, *uczucie* 5} 'emotion', which is linked by hypernymy with 236 other sets of synonyms (however, many of them include only one lexical unit).

We assumed that hyponyms, direct and indirect, share the emotional aspect of their meanings with their hypernyms. Following this assumption, in order to expand the set of lemmas developed on the basis of Markowski's list of the lexis common for different genres (Markowski, 1992), we performed analysis of lemmas from this list from the perspective of expressing emotions. We tried to identify those lemmas[5] that express clear emotional polarity: positive or negative. Next, for each selected lemma we identified all lexical units from plWordNet that correspond to it. Starting with the identified lexical units and the sets of synonyms they belong to — called *synsets*, i.e. places or nodes in the plWordNet graph, we searched for hypernymic subtrees, such that

- they included as many of the identified synsets as possible,
- they encompass lexical units of the same emotional polarity,
- and they are as large as possible, in terms on the number of synsets included.

The initial subtrees delineated at the beginning were quite large and rooted in more general lexical units as hypernyms. However, they were next dived into smaller ones (defined by hypernyms dominating smaller number of lexical units), but more coherent from the point of view of expressing emotions or emotional polarity of their lexical units. We have also explored possibilities of expanding the initial subtree originating from the meaning of a lemma from Markowski's list of the lexis common for different genres (Markowski, 1992) by moving the subtree root to more general hypernyms. However, such a generalising expansion mostly resulted in too far going enlargement of the subtree and lost of coherence with respect to the emotions expressed. In this work we used the version 2.0 of plWordNet, as the version 2.2 was not yet available in the moment of expanding the list.

4. Names of emotions, affects and emotional states
As the first subset of words, we analysed lemmas that name emotions, but they are not emotionally polarised (Spagińska-Pruszak, 1994, p. 11). We do not equal

---

[5]In the case of each from the list, we analysed all lexical units represented by it, as one lemma can correspond to several lexical units.

{ miłość 1, uczucie 4, afekt 1 } 'love' hypernymy (5/24)

   { miłość 1, uczucie 4, afekt 1 } is a hypernym of { namiętność 3 } '≈love'

   { miłość 1, uczucie 4, afekt 1 } is a hypernym of { zakochanie 1, zadurzenie 1 } '≈falling in love'

   { miłość 1, uczucie 4, afekt 1 } is a hypernym of { uwielbienie 1, adoracja 1 } 'worship, adoration'

      { uwielbienie 1, adoracja 1 } is a hypernym of { ubóstwienie 2 } '≈ worship, adoration'

      { uwielbienie 1, adoracja 1 } is a hypernym of {fetyszyzacja 1, fetyszyzm 1 } '≈ idealisation, glorification'

      { uwielbienie 1, adoracja 1 } is a hypernym of { narcyzm 1, samouwielbienie 1, autofilia 1, idolatria 1, samozachwyt 1 }

         { narcyzm 1, samouwielbienie 1, autofilia 1, idolatria 1, samozachwyt 1 } 'self-love, narcism' is a hypernym of { próżność 1 } '≈ conceit, conceitedness, vanity'

            { próżność 1 } is a hypernym of { snobizm 1 } 'snobbery, snobbism, snobbishness'

               { snobizm 1 } is a hypernym of { arystokratyzm 1 } '≈ snobbery, snobbism, snobbishness'

      { uwielbienie 1, adoracja 1 } is a hypernym of { kult 1 } '≈ worship, adoration'

         { kult 1 } is a hypernym of { kult jednostki 1 } 'cult of personality'

      { uwielbienie 1, adoracja 1 } is a hypernym of { bałwochwalstwo 1 } '≈ worship, adoration'

      { uwielbienie 1, adoracja 1 } is a hypernym of { cielęcy zachwyt 1, cielęce uwielbienie 1 } '≈ worship, adoration'

   { miłość 1, uczucie 4, afekt 1 } is a hypernym of { miłość 2 } 'agape, agape love'

      { miłość 2 } is a hypernym of { patriotyzm 1, miłość do ojczyzny 1 } 'patriotism, nationalism'

   ...

Figure 1: An example of a hypernymic subtree rooted in the synset (a set of near synonyms) { miłość 1, uczucie 4, afekt 1 } 'love' from plWordNet 2.0.

feeling of an emotion with a way of expressing it. We have been interested only in those emotions that were expressed in the text by the lexical means. We decided that the presence of emotions in a text and the way of expressing them can be a feature which characterises the speaker and the text topics, as well as it can help us to distinguish between genuine and counterfeited suicide notes. Following A.

Spagińska-Pruszak we did not apply psychological theories on separating emotions and affects, so we treated both notions as synonymous (Spagińska-Pruszak, 1994, p. 29–42).

We started search for the names of emotions from a list of nouns and verbs selected from Markowski's list of the lexis common for different genres.

4.1. Nominal names of emotions

The group of nouns from the 9th field of (Markowski, 1992, pp. 115–116): *Uczucia, emocje, oceny uczuciowe i emocjonale — Subpole 'uczucia'* (Affects, emotions, emotional assessments and emotionals — The Subfield 'affects') has been expanded with lexical units from the plWordNet 2.0 hypernymic subtree dominated by the synset: {*emocje* 1, *uczucie* 5} 'emotion'. As a result, we received a list of lexical units of 272 different lemmas. Among them, 66 lemmas i.e. 24,26% of the list, occur in PCSN. In the case of all those lemmas, but one, namely *obrzydzenie* 'disgust', if a lemma occurs in SCSN, it also occurs in PCSN. Only 21 lemmas (7,72%) occur only in SCNS. Concerning the lemma frequency in PCSN, the top five positions are: *miłość* 'love', *nadzieja* 'hope', *serce* 'heart', *żal* 'regret', *szczęście* 'happiness'. The authors of the counterfeited suicide notes used most often the lemma *nadzieja* 'hope': it occurs in 6,84% of documents in PCSN and in 19,66% documents of SCSN.

4.2. Verbal names of emotions

Verbs used as the lexical markers of emotions are based on the group of 92 verbs from the 9th field of (Markowski, 1992, pp. 115–116): *Uczucia, emocje, oceny uczuciowe i emocjonale — Subpole 'uczucia'* (Affects, emotions, emotional assessments and emotionals — The Subfield 'affects') that has been next expanded with lemmas from a few identified hypernymic verb subtrees from plWordNet. Contrary to (Markowski, 1992), verbs from aspectual pairs have been treated as separate lexical units.[6] In addition, we have also expanded the verb set with morphological derivatives of the verbs already included in the set that have been found in PCSN and SCSN merged together, e.g.: *kochać* 'to love' — *ukochać* 'to love$_{perf.}$', *odkochać* 'to stop to love'; *czuć* 'to feel' — *uczuć* 'to feel$_{perf}$', *cierpieć* 'to suffer' — *nacierpieć się* 'to suffer$_{perf.,iteratively}$', *nienawidzić* 'to hate' — *znienawidzić* 'to hate$_{perf.}$', *szanować* 'to respect' — *uszanować* 'to respect$_{perf.}$'. 46 verb lemmas, i.e. 50% from the created set have been found on the PCSN frequency list, and 22 (23.91%) have been found on the SCSN frequency list. The top frequent verb lemmas found on the PCSN list are: *kochać* 'to love', *czuć (się)*,[7] *bać się* 'to be afraid', *żałować* 'to regret', *nienawidzić* 'to hate'. The first top positions on the SCSN verb frequency list are occupied by: *kochać*, *czuć (się)*, *bać się*, *cierpieć* 'to suffer', *martwić się*. The first verb, i.e. *kochać* 'to love' dominates in both the genuine and counterfeited notes — it can be found in 41.4% genuine documents and in 36.75% counterfeited documents. The second verb according to the rank, i.e. *czuć (się)* 'to feel, to feel

---

[6]This is in accordance to the solution assumed in plWordNet which was motivated by different lexico-semantic relations that both elements of the aspectual pair participate in, cf Maziarz et al. (2013).

[7]Due to limitations of the automated lemmatisation it is hard to distinguish between *czuć* 'to feel' and *czuć się* 'to feel like'.

like' have slightly larger frequency in SCSN 11.11% than in PCSN 6.19%. It seems that the use of this verb allows for speaking about emotions in a less direct way, e.g.

Jestem samotny. 'I am lonely'

vs

Czuję się samotny. 'I feel lonely'

In the first case the speaker names his emotional state, while in the second he is analysing his emotional state.

The third lemma in the rank list list, i.e. *bać się* 'to be afraid' occurred in about 4% of genuine and counterfeited notes. The similar frequencies are expressed by other verbs among the five most frequent verbs of SCNS: *cierpieć* 'to suffer' and *martwić się* 'to worry', as well as *żałować* 'to regret' and *nienawidzić* 'to hate' that appear relatively frequently in PCSN.

4.3. Names of emotions — conclusions

On the basis of the analysis of names of emotions, affects and emotional states expressed with the help of nouns and verbs, we can claim that:

1. Authors of the genuine suicide notes more often write about emotions than the authors of the counterfeited notes.

2. In PCSN the most frequently used lemmas thematically related to emotions are the verb *kochać* 'to love' and the nouns: *miłość* 'love', *nadzieja* 'hope' and *serce* 'heart'.

3. In SCSN, the verb *kochać* 'to love' is also the most frequently used verb thematically related to emotions, but the most frequent noun related to emotions is *nadzieja* 'hope', which occurs much more often in the counterfeited notes (including a collocation *mieć nadzieję* '≈to hope') than in the genuine suicide notes. Moreover, the noun *nadzieja* 'hope' is mostly used in the present tense in the counterfeited notes (only one past form of *mieć nadzieję* '≈to hope' has been found), while this noun is mostly used with the reference to the past and to the presence in the genuine notes, e.g. *miałem nadzieję* 'I had hope, hoped, *żyłem nadzieją* 'I lived with the hope', *wypaliłeś we mnie malutką nadzieję* 'You have burnt the smallest hope out of me'.

5. Emotion operators

Emotions can be expressed not only explicitly, but also with the means of *emotion operators* that are dependent units (lexical or phrasal) used to modify independent units, i.e. the emotion operators do not have meaning by their own, but they are used to modify the meaning of other units, e.g. the operators are used to express hierarchical relations between independent units (Lewiński, 2006, pp. 54–55). Emotion operators can be divided into three groups:

1. exclamations encompassing primary exclamation words, parenthetical exclamation words (Grochowski, 1997, p. 14) and affective modifiers (Jodłowski, 1976, p. 21);

2. emotive phrases with positive polarisation and phrases "expressing emotions without any specified positive or negative profile" (Rodak, 2000, p. 196) e.g. *na szczęście* 'luckily', *dzięki Bogu* 'thanks God';

3. generalising operators: generalising pronouns and negative pronouns, e.g. *wszystko* 'everything', *żaden* 'none', *nikt* 'nobody', *zawsze* 'always'.

5.1. Exclamation words

The most frequent examples of exclamation words that can be found in the suicide notes are affective modifiers, e.g. *nawet* 'even' or *raczej* 'rather'. Parenthetical exclamations are much more rare, as curses are not characteristic words for the analysed texts, e.g. *cholera* 'damn' occurs 0.24% in PCSN and *kurwa* 'shit, fuck' 1.59% in PCSN. We could find only occasionally primary exclamation words in the genuine notes from PCSN, e.g. *ach* 0.12%, *oj* 0.12%, *hehe* 0.49%. They are more characteristic for spoken texts.

Affective modifiers occur in both corpora (namely PCSN and SCSN) with the similar frequency. Among them, the highest positions on the rank list are occupied by *nawet* 'even' and *przecież* 'but, yet, though'. However the first is much more frequent: *nawet* 9.88% in PCSN, *przecież* 3.90% in PCSN; *nawet* 11.11% in SCSN, *przecież* 2.56% in SCSN. Larger variety of affective modifiers is a characteristic feature of the genuine suicide notes. We could not find lemmas: *nareszcie* 'at last', *koniecznie* 'necessarily', *bynajmniej* 'not at all, not in the least', *na szczęście* 'luckily', *raczej* 'rather' in the counterfeited notes. This is definitely caused by the fact that those notes have been written by the smaller number of different authors. However, it can only be a signal that those utterances are more schematic.

5.2. Emotive phrases

The set of emotive phrases found in both analysed corpora is dominated by phrases of the religious character, e.g. *mój Boże!* 'my God!', *Boże!* 'God!' or *Jezu!* 'Jesus!'. However, mostly, they are not emotion operators in suicide notes, but genuine address forms related to prayers. As there is no content related to religion in the counterfeited notes, there are no religious address forms at all in those documents.

5.3. Generalising operators

Generalising operators occur in both the genuine and counterfeited suicide notes. Several generalising pronouns can be found in both corpora:

- in PCSN: *wszystko* 'everything', *zawsze* 'always', *wszyscy* 'all (people)', *każdy* 'every, everyone', *wszędzie* 'everywhere';

- and in SCNS: *wszystko* 'all (things)' , *zawsze* 'always', *wszyscy* 'all (people)', *każdy* 'every, everyone'.

In a similar way, some negative pronouns can be also noticed in both corpora:

- in PCSN: *nic* 'nothing', *nikt* 'nobody', *nigdy* 'never', *żaden* 'none', *niczyj* 'of nobody',

- in SCNS: *nic* 'nothing', *nikt* 'nobody', *nigdy* 'never', *żaden* 'none'.

It is worth to emphasised that the negative pronoun *nic* 'nothing' has much larger frequency in SCSN: 29.91% than in PCSN: 15.15%. This imbalance is probably caused by the situation of the writer in the moment of counterfeiting the note. The writer has little or no knowledge about the described events and has no other choice but to use generalising operators like 'nothing' etc. The use of the generalising operator simplifies the creation of the text in a moment in which the amount of knowledge possessed by the write about the described events is very small. The comparison of the two frequency lists of PCSN i SCSN shows that the authors of the counterfeited notes less frequently use concrete words, and much more frequently abstract words like *problem* 'problem', *sens* 'sens' or *sytuacja* 'situation'. The use of generalising operators also allows for avoiding informing about concrete events and persons.

5.4. Emotion intensifiers
The class of emotion intensifiers has been defined by Dagmara Bałabaniak (2007, p. 15) as

> "lexical units that introduce the meaning of the high level of a feature which characterises the given object in an absolute way. Lexical emotion intensifier always occurs in relation with some other predicate; it modifies a features named by this predicate (it does introduce any features by itself)."

For instance, the word *bardzo* 'very' and its synonyms are unequivocally examples of emotion intensifiers, as it was described in the work (Janus, 1977, p. 5). The list of *bardzo* synonyms encompasses *diablo* 'bloodily', *diabelnie* 'bloodily'/*diabelny* 'bloody', *cholernie* 'damn'/*cholerny* 'damn', *kolosalnie* 'enormously'/*kolosalny* 'enormous', *mocno* 'very much'/*mocny* 'strong', *ogromnie* 'hugely/*ogromny* 'huge', *olbrzymi* 'gigantic', *okropnie* 'terribly'/*okropny* 'terrible', *piekielnie* 'terribly'/*piekielny* 'terrible', *potwornie* 'horribly'/*potworny* 'horrible', *silnie* 'strongly'/*silny* 'strong', *strasznie* 'terribly, awfully'/*straszny* 'terrible', *straszliwie* 'horribly'/*straszliwy* 'horrible', *szalenie* 'extremely, madly'/*szalony* 'extreme, mad', *wielce* 'very much', *wielki* 'great'/*duży* 'big, large', *wysoce* 'highly'/*wysoko* 'high' (Janus, 1977, pp. 43–44). As E. Janus concentrated in her work on the synonyms of the word *bardzo*, the above list must be complemented. Thus, for the needs of our analysis, we have expanded the list of emotion intensifiers with:

- the word *bardzo* 'very' and its synonyms (Janus, 1977) — in PCSN: *bardzo* 'very', *mocno* 'strongly', *wielki* 'great', *duży* 'big', *strasznie* 'terribly'; in SCSN: *bardzo* 'very', *mocno* 'strongly', *okropny* 'terrible', *wielki* 'great', *duży* 'big, large';
- words representing the limit of a feature — they are rare in both corpora — in PCSN: *zupełnie* 'completely', *całkiem* 'completely, entirely', *całkowicie* 'completely', *ostatecznie* 'finally, defnitely', *totalnie* 'totally'; in SCSN: *ostatecznie* 'finally';
- words representing the lower level of a feature — in PCSN *mało* 'little', *dość* 'enough', *dosyć* 'enough', *niewiele* 'not much, little', *nieco* 'somewhat'; in SCNS: *dość* 'enough', *dosyć* 'enough', *mało* 'litle', *niewiele* 'not much, little',

- and also other words expressing high level of a feature.

The words of the last group are almost absent in PCSN — we can find only *dobrze* 'well' in the sense *mocno* 'strongly', *sporo* 'a lot, much', *wyjątkowo* 'exceptionally', and they do not occur in SCNS at all.

Neutral emotion intensifier *bardzo* 'very' occur in 29.15% of genuine notes and in 17.95% of counterfeited notes. Collocations of intensifiers with *kochać* 'to love' are among the most frequent collocations in both corpora, e.g.

    Kocham Was bardzo. 'I love you very much'

Other frequent collocations including the intensifier *bardzo* 'very, very much' are, e.g.:

bardzo przepraszać 'to appologise a lot' in *Przeproś Mamusię moją bardzo.*, 'Appologise my mother very much'
    bardzo boleć 'to suffer a lot' in *Żegnam! / Choć mnie to bardzo boli.* 'Farewell! However, it hurts me very much',

bardzo chcieć 'to want very much' in *Tak bardzo chciałbym żyć.* 'I wanted to live so much',

bardzo odmienić 'to change a lot' in *nie wiesz jak bardzo odmieniłeś moje życie* 'you do not know how much you have changed my life',

bardzo starać się 'to try very much' in *Tak bardzo się / Starałem* 'I have been trying so much',

bardzo brakować 'to miss a lot' in *Mamo tak bardzo mi Ciebie brakuje!* 'Mummy, I miss you so much!',

bardzo prosić 'to ask very much' in *A tak bardzo Cię prosiłem* 'I have been asking you so much',

bardzo skrzywdzić 'to harm a lot' in *Skrzywdziłam Was bardzo* 'I have harmed you very much',

bardzo szkoda 'big pity' in *Bardzo mi szkoda/ moich kochanych dzieci* '≈I am so sorry about my dear children',

bardzo płakać 'to cry very much' in *Wtedy bardzo płakałem, ale nikt nie widział* 'I was crying a lot on that time, but nobody saw it',

bardzo cierpieć 'to suffer very much' in *Proszę tylko nie cierp tak bardzo* 'Please, only do not suffer so much',

bardzo żałować 'to be very sorry' in *Żałuję bardzo, że Cię nie/ posłuchałam.* 'I am very sorry that I have not listen to you',

bardzo błagać 'to beg a lot' in *Zrozumcie mnie bardzo was błagam.* 'Please, understand me, I am begging you very much',

bardzo się bać 'to be afraid a lot' in *bardzo się boję!* 'I am very afraid'.

Among adjective collocations with the intensifier *bardzo* 'very' we can find:

bardzo dobry 'very good' in *byłaś dla mnie bardzo dobra* 'you have been very good to me',

bardzo duży 'very big' in *Zrobiliście mi bardzo dużą krzywdę* 'You have done a lot of harm to me',

bardzo zmęczony 'very tired' in *Jestem już bardzo zmęczony* 'I am very tired',

bardzo serdecznie 'very heartilly/cordially' in *Bardzo serdecznie Księdza przepraszam* ''I appologise very heartilly you, padre',

bardzo uczuciowa 'very passionate' in *JESTEŚ BARDZO UCZUCIOWA* 'You are very passionate',

bardzo bolesne 'very painful, sore' in *Było to dla mnie bardzo bolesne* 'It was very painful/sore for me',

bardzo chory 'very ill' in *Byłam bardzo chora* 'I have been very ill',

bardzo szczęśliwy 'very happy' in *Teraz jesteś już wolna i/pewnie bardzo szczęśliwa* 'You are now free already and / very happy, I suppose'.

We have also found a number of adverbial collocations, e.g.:

bardzo przykro 'very sorry' in *Bardzo jest mi / przykro za to co /zrobiłem* 'I am very / sorry for what / I have done',

bardzo dużo 'very much', *bardzo ciężko* 'very heavily' in *i wierzcie mi że bardzo ciężko /jest z tym żyć* 'and believe me, it is very hard / to live with this',

bardzo długo 'very long' in *Już bardzo długo nie miałam takich myśli.* 'For a long time I have not had such thoughts',

bardzo źle 'very bad' in *Było mi bardzo źle* 'I felt very bad',

bardzo niedobrze 'very improperly, wrongly' in *Niestety już od dłuższego czasu działo się ze mną/ bardzo niedobrze* 'Unfortunately, it has been wrong with me for a long time / very wrong',

bardzo dobrze 'very good' in *i to bardzo dobrze* 'and very good'.

The word *mocno* 'heavily/strongly' is the second most frequent emotion intensifier in both corpora, but it is much rarer in the analysed texts: *bardzo* 'very' occurs in 29.15% of PCSN documents, while *mocno* 'strongly' in 5.37%, in SCSN *bardzo* 'very' is present in 17.95% of documents, while *mocno* 'strongly' in 4.27%. Moreover, this lemma mostly occurs only in the greeting formula of the suicide note, e.g.:

Całuję Cię mocno 'I am kissing you very much',

Kocham Cię mocno 'I am loving you very much',

Mocno pozdrawiam 'I am greeting you very much'.

Besides the dominating collocation with the verb *kochać* 'to love', we can also find such associations as

Chcę Cię mocno przeprosić 'I would like to apologise you very much',

bardzo mocno mnie zdradziły 'they were very much unfaithful to me', *zbyt mocno ufałem ludziom* 'I trusted people too much'.

6. Valuation words

We collected the initial set of valuation words from the subfield *Oceny* (Assessments) which belongs to several fields of Markowski's list of the lexis common for different genres:

- *Uczucia, emocje, oceny uczuciowe i emocjonalne*; 'Affects, emotions, emotional assessments'
- and *Rozum, pamięć, wyobraźnia i oceny z nimi związane* 'Mind, memory, imagination and assessments related to them'.

We added also to the list words from the subfield *Cechy i oceny zachowania i postępowania* 'Features and evaluation of the behaviour and conduct' from the fields of (Markowski, 1992):

- *Charakter człowieka i jego cechy* 'Human character and its features';
- *Moralność i oceny z nią związane* 'Morality and assessments related to it';
- *Zachowanie i postępowanie* 'Behaviour and conduct'.

The initial list was next expanded on the basis of plWordNet 2.0 in a way following the procedure described in the previous sections. The obtained list goes beyond the limits of Markowski's list of the lexis common for different genres. As a result we have obtained a list of 488 adjectives. 151 of them (i.e. 30.94% of the whole list) were found in PCSN and only 41 in SCSN (8.40% of the list.). The five adjectives from the list with the highest frequencies in PCSN are: *kochany*, *dobry*, *zły*, *szczęśliwy* and *nowy* 'new'. The last one occurs in two meanings: *nowy* 1 'new — describing something that will be created or happen', e.g. in *nowa sytuacja* 'new situation' and *nowy* 2 'new — characterising something that have just started to exist', e.g. in *nowe życie* 'new life'. The most frequent adjectives in SCSN are: *kochany* 'dear, beloved', *dobry* 'good', *drogi* 'dear', *zły* 'bad', *jedyny* 'only, sole, the one and only one', e.g. *jedyna osoba* 'the only person', *jedyna kobieta* 'only woman', *jedyni bliscy* 'the only close relatives'. Some of them come from the address forms that is especially visible in the case of the adjectives: *kochany* 'dear' and *drogi* 'dear'. However, as it was shown above, the biggest difference between both corpora, concerning this type of adjectives is in the different numbers of adjectives found in both corpora.

7. Terms of endearment and invectives

The closest relation between the speaker and receiver of the suicide note is expressed with the help of terms of endearment, i.e. appellatives occurring "in the form of words or expressions used in the situation of especial intimacy, mostly [...] in relationships between spouses, fiancés, lovers or in the relation parents — children" (Perlin & Milewska, 2000, p. 165). The terms of endearment are also defined as "intimate nicknames that people give to their life partners, family members or close friends" (Bańko & Zygmunt, 2010, p. 6). 120 terms of endearment were found in PCSN and can be grouped into the following classes:

- animal names, e.g. *Miśku* 'Teddy Bear', *Mój Misiu* 'My Teddy Bear', *Misiuniu* '(small) Teddy Bear', *Misiaczku* '(my small) Teddy Bear', *Żabciu* '(my small) Frog', *Rybko* '(my small) Fish', *Myszko* '(my small) Mouse', *Kurczaczku* '(my little) Chicken';
- lexemes of the meaning 'happiness, dear/beloved' that are also described as "names expressing relation between the speaker and the addressee" (Bańko & Zygmunt, 2010, p. 130), e.g.: *Kochanie* 'My Dear', *Kochana* 'My Dear', *Kochani* 'My Dear', *Kochani Wy Moi* 'You, My Dear', *Ukochana moja* 'my beloved';
- lexemes of the meaning 'something precious' e.g.: *Skarbie* 'My precious', *Skarbeńku* 'my precious (diminutive)', *Jedyna* 'My Dear';
- tender names for small children, e.g.: *Niuniu* '(tenderly) small girl', *Niuńko* '(tenderly) small girl';
- lexemes related to astronomy or meteorology e.g.: *Słonko* '(diminutive) sun';
- names of the family roles e.g.: *żoneczko* '(diminutive) wife', *tatusiu* "(diminutive) daddy' .

In SCSN we could find only 22 terms of endearment and most of them belong only to the two classes: lexemes of the meaning 'happiness, dear/beloved', e.g. *Kochanie* 'my dear', *Kochana* 'my dear (female)', *Kochany* 'my dear (male)' and lexemes of the meaning 'something precious', e.g. *Drogi* 'dear', *Najdroższa* 'my most precious'.

In the genuine notes terms of endearment occur not only in the expressions directed towards the addressee (address forms), in a characteristic vocative case (noun phrases), but they can be found also in the note content or signature, e.g. *Twój Kotek* 'yours kitty'. Such phrases are traces of the intimate communication between the speaker and the addressee from the time before the moment of writing the note, e.g.

Jesteś moim Misiem (. . .) *Kocham Cię Misiaczku* 'You are my Teddy Bear (...) I love you Teddy Bear'.

In SCSN, we can find members of only one of the above classes, namely words from the class of lexemes of the meaning 'happiness, dear/beloved', e.g. *Kochanie* 'My Dear'.[8]

---

[8]In general, *Kochanie* 'my dear' is the most frequent Polish term of endearment (cf Bańko & Zygmunt, 2010, p. 124).

Invectives appear relatively rarely in the suicide notes, e.g.

ty ochydna kłamco 'you awful liar',

Ty IDIOTKO 'you idiot',

Twoja zasługa /jest w tym Ty babiożu 'yours merit / is in that you awful woman (negative augmentative)'.

This group of words is associated with language aggression and it is characteristic for the genuine suicide notes in which the occurrence of invectives is motivated not only by the life situation but also by the acquaintance with the addressee. That enables maximal shortening the communication distance to the addressee. Invectives do not occur at all in SCSN.

8. Conclusions and future work

Emotions, affects and emotion valuations appear significantly more often as topics of the genuine suicide notes than the counterfeited notes. In PCSN we can observe much larger variety of the vocabulary related to emotions in terms of both: means of expressing emotions and the use of different synonyms. This is related to the fact that the group of different writers is much bigger in the case of PCSN than in the case of SCSN, and also the lack of the knowledge among the authors in SCSN about the context of the described events. Moreover, in the case of SCSN there is no common knowledge (common communication context) shared between the speaker and the addressee. Thus the difference between both corpora originate from the influence of the original situation on the process of writing the text.

References to love are present in the majority of texts in both corpora, e.g. *miłość* 'love', *serce* 'heart', *kochać* 'to love', *kochany* 'dear, beloved', and *kochanie* 'dear, beloved', ... However, the counterfeited notes differ in the use of the word *nadzieja* 'hope'. It mostly refer to the presence in these notes. PCSN also differentiates from SCSN in the higher frequency of the intensifier *bardzo* 'very'. It is worth to notice, that the frequency of different lemmas that are identical or very similar in PCSN and SCSN may be caused by the different motivation for their use. Thus they may be used in slightly different meaning in both corpora. For instance, the generalising operators have similar ranks in PCSN and SCSN, but they are used to emphasise commonality of events and their consistent character. However, in SCSN they express the lack of knowledge or limited knowledge of the speaker about the context of the described situation.

References

Bałabaniak, D. (2007). Semantyczne ograniczenia intensyfikatorów leksykalnych. *Prace Filologiczne, 52*, 15–22.

Bańko, M. & Zygmunt, A. (2010). *Czułe słówka: słownik afektonimów.* Warszawa: PWN.

Grochowski, M. (1997). *Wyrażenia funkcyjne: Studium leksykalne.* Kraków: PAN, Instytut Języka Polskiego.

Grochowski, M. (2009). Intensyfikatory, partykuły a negacja: O homonimii gramatycznej dwóch jednostek monosemicznych o postaci w ogóle. In K. Waszakowa, T. Korpysz, & J. Chojak (Eds.), *Człowiek — słowo — świat* (pp. 306–315). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Grzegorczykowa, R. (1981). *Wykładniki intensywności cechy (na materiale polskim i rosyjskim)*. Wrocław: Zakład Narodowy imienia Ossolińskich.

Grzegorczykowa, R. (1995). *Wprowadzenie do semantyki językoznawczej*. Warszawa: PWN.

Janus, E. (1977). Wykładniki intensywności cechy we współczesnej polszczyźnie. *Polonica*, *3*, 47–57.

Jodłowski, S. (1976). *Podstawy polskiej składni*. Warszawa: PWN.

Lewiński, P. (2006). Operatory emocji. In K. Michalewski (Ed.), *Wyrażanie emocji* (pp. 54–55). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Markowski, A. (1992). *Leksyka wspólna różnym odmianom polszczyzny* (Vols. 1–2). Wrocław: Wydawnictwo „Wiedza o kulturze".

Maziarz, M., Piasecki, M., & Szpakowicz, S. (2012). Approaching plWordNet 2.0. In C. Fellbaum & P. Vossen (Eds.), *Proceedings of 6th International Global Wordnet Conference*, The Global WordNet Association (pp. 189–196).

Maziarz, M., Piasecki, M., & Szpakowicz, S. (2013). The chicken-and-egg problem in wordnet design: Synonymy, synsets and constitutive relations. *Language Resources and Evaluation*, *47*(3), 769–796. `http://doi.org/10.1007/s10579-012-9209-9`

Styrcz-Przebinda, L. (1997). Parę uwag o możliwościach klasyfikacji wykrzykników. *Polonica*, *18*, 163–167.

Perlin, J. & Milewska, M. (2000). Afektonimy w polskim, francuskim, hiszpańskim i niderlandzkim: Analiza morfologiczna i semantyczna. In I. Nowakowska-Kempna, A. Dąbrowska, & J. Anusiewicz (Eds.), *Uczucia w języku i tekście* (pp. 165–173). Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego. (*Język a kultura*, 14)

Rodak, R. (2000). Frazemy jako emotywne operatory interakcyjne. In I. Nowakowska-Kempna, A. Dąbrowska, & J. Anusiewicz (Eds.), *Uczucia w języku i tekście* (pp. 187–198). Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego. (*Język a kultura*, 14).

Spagińska-Pruszak, A. (1994). *Język emocji: Studium leksykalno-semantyczne rzeczownika w języku polskim, rosyjskim i serbsko-chorwackim*. Gdańsk: Wyd. Uniwersytetu Gdańskiego.

Tuckman, R., Kleiner, R. J., & Lavell, M. (1959, July). Emotional content of suicide notes. American Journal of Psychiatry, 116(1), 59–63. `http://doi.org/10.1176/ajp.116.1.59`

Zaśko-Zielińska, M. (2013). *Listy pożegnalne: W poszukiwaniu wyznaczników autentyczności tekstu*. Wrocław: Quaestio.

The authors declare that they have no competing interests.

The authors' contribution was as follows: concept of the study: Monika Zaśko-Zielińska; data analyses: Monika Zaśko-Zielińska and Maciej Piasecki; the writing: Monika Zaśko-Zielińska and Maciej Piasecki.